

Whitaker, Jerry C. “Chapter 10 – Device Performance Criteria”
Power Vacuum Tubes Handbook 2nd Edition.
Ed. Jerry C. Whitaker
Boca Raton: CRC Press LLC, 2000

Device Performance Criteria

10.1 Introduction

Many shortcomings in the received audio, video, and/or data stream of a transmission resulting from noise, reflections, and terrain limitations cannot be controlled by the equipment design engineer or the operating maintenance engineer. However, other problems can be prevented through proper application of vacuum tube devices and proper equipment adjustment. The critical performance parameters for a transmission system include:

- Output power
- Transmission line/antenna VSWR
- Frequency stability/accuracy
- Occupied bandwidth
- Distortion mechanisms, including *total harmonic distortion (THD)*, *intercarrier phase distortion*, and *quadrature distortion*
- Differential phase and gain
- Frequency response errors
- Envelope delay

The adjustment schedule required to maintain a satisfactory level of performance depends upon a number of factors, including:

- Operating environment
- Equipment quality
- Age of the equipment
- Type of service being performed

10.2 Measurement Parameters

Most measurements in the RF field involve characterizing fundamental parameters. These include signal level, phase, and frequency. Most other tests consist of measuring these fundamental parameters and displaying the results in combination by using some convenient format. For example, signal-to-noise ratio (S/N) consists of a pair of level measurements made under different conditions and expressed as a logarithmic ratio. Modern test instruments perform the necessary manipulation of the measured parameters to yield a direct readout of the quantity of interest.

When characterizing a vacuum tube device or stage, it is common to view it as a box with input and output terminals. In normal use, a signal is applied to the input of the box, and the signal—modified in some way—appears at the output. Some measurements are *one-port* tests, such as impedance or noise level, and are not concerned with input/output signals, but with only one or the other.

Measurements are made on equipment to check performance under specified conditions and to assess suitability for use in a particular application. The measurements may be used to verify specified system performance or as a way of comparing several pieces of equipment for use in a system. Measurements also may be used to identify components in need of adjustment or repair.

Many parameters are important in vacuum tube devices and merit attention in the measurement process. Some common measurements include frequency response, gain or loss, harmonic distortion, intermodulation distortion, noise level, phase response, and transient response.

Measurement of level is fundamental to most RF specifications. Level can be measured either in absolute terms or in relative terms. Power output is an example of an absolute level measurement; it does not require any reference. S/N and gain or loss are examples of relative, or ratio, measurements; the result is expressed as a ratio of two measurements. Although it may not appear so at first, frequency response is also a relative measurement. It expresses the gain of the device under test as a function of frequency, with the midband gain (typically) as a reference.

Distortion measurements are a way of quantifying the unwanted components added to a signal by a piece of equipment. The most common technique is total harmonic distortion, but others are often used. Distortion measurements express the unwanted signal components relative to the desired signal, usually as a percentage or decibel value. This is another example of multiple level measurements that are combined to give a new measurement figure.

10.2.1 Power Measurements

The simplest definition of a level measurement is the alternating current amplitude at a particular place in the system under test. However, in contrast to direct current measurements, there are many ways of specifying ac voltage in a circuit. The most common methods include:

- Average response

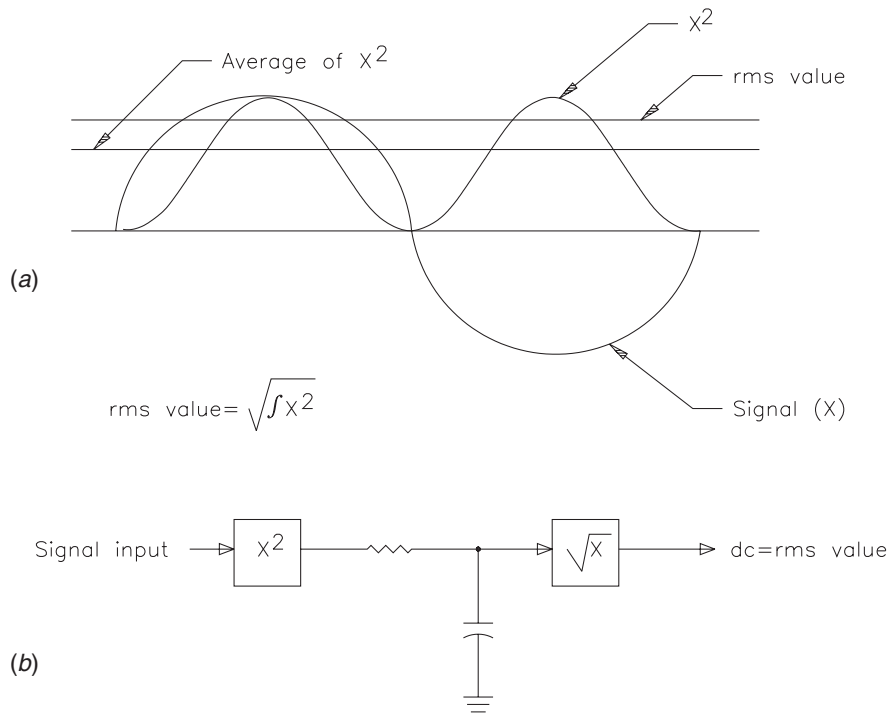


Figure 10.1 Root-mean-square (rms) voltage measurements: (a) the relationship of rms and average values, (b) the rms measurement circuit.

- Root mean square (rms)
- Peak

Strictly speaking, the term *level* refers to a logarithmic, or decibel, measurement. However, common parlance employs the term for an ac amplitude measurement, and that convention will be followed in this chapter.

Root Mean Square

The root-mean-square technique measures the effective power of the ac signal. It specifies the value of the dc equivalent that would dissipate the same power if either were applied to a load resistor. This process is illustrated in Figure 10.1 for voltage measurements. The input signal is squared, and the average value is found. This is equivalent to finding the average power. The square root of this value is taken to transfer the signal from a power value back to a voltage. For the case of a sine wave, the rms value is 0.707 of its maximum value.

Assume that the signal is no longer a sine wave but rather a sine wave and several harmonics. If the rms amplitude of each harmonic is measured individually and added,

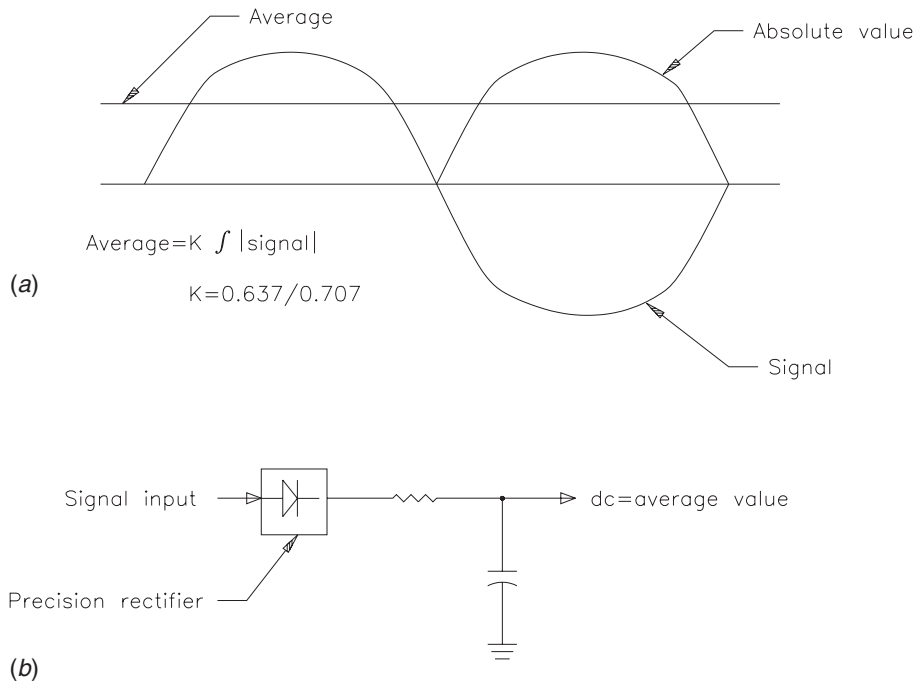


Figure 10.2 Average voltage measurements: (a) illustration of average detection, (b) average measurement circuit.

the resulting value will be the same as an rms measurement of the signals together. Because rms voltages cannot be added directly, it is necessary to perform an rms addition. Each voltage is squared, and the squared values are added as follows:

$$V_{rms\ total} = \sqrt{V_{rms\ 1}^2 + V_{rms\ 2}^2 + \dots + V_{rms\ n}^2} \quad (10.1)$$

Note that the result is not dependent on the phase relationship of the signal and its harmonics. The rms value is determined completely by the amplitude of the components. This mathematical predictability is useful in practical applications of level measurement, enabling correlation of measurements made at different places in a system. It is also important in correlating measurements with theoretical calculations.

Average-Response Measurement

The average-responding voltmeter measures ac voltage by rectifying it and filtering the resulting waveform to its average value, as shown in Figure 10.2. This results in a dc voltage that can be read on a standard dc voltmeter. As shown in the figure, the average value of a sine wave is 0.637 of its maximum amplitude. Average-responding

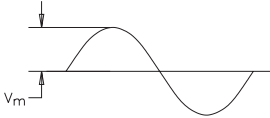
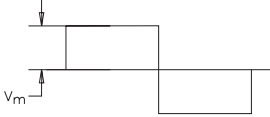
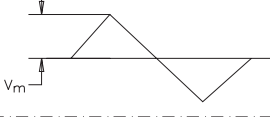
Waveform	rms	Average	rms average	Crest factor
 Sine wave	$\frac{V_m}{\sqrt{2}}$ $0.707 V_m$	$\frac{2}{\pi} V_m$ $0.637 V_m$	$\frac{\pi}{2\sqrt{2}} = 1.111$	$\sqrt{2} = 1.414$
 Square wave	V_m	V_m	1	1
 Triangular wave or sawtooth wave	$\frac{V_m}{\sqrt{3}}$	$\frac{V_m}{2}$	$\frac{2}{\sqrt{3}} = 1.155$	$\sqrt{3} = 1.732$

Figure 10.3 Comparison of rms and average voltage characteristics.

meters usually are calibrated to read the same as an rms meter for the case of a single sine wave signal. This results in the measurement being scaled by a constant K of $0.707/0.637$, or 1.11 . Meters of this type are called *average-responding, rms calibrated*. For signals other than sine waves, the response will be different and difficult to predict.

If multiple sine waves are applied, the reading will depend on the phase shift between the components and will no longer match the rms measurement. A comparison of rms and average-response measurements is made in [Figure 10.3](#) for various waveforms. If the average readings are adjusted as described previously to make the average and rms values equal for a sine wave, all the numbers in the *average* column would be increased by 11.1 percent, and the *rms-average* numbers would be reduced by 11.1 percent.

Peak-Response Measurement

Peak-responding meters measure the maximum value that the ac signal reaches as a function of time. This approach is illustrated in [Figure 10.4](#). The signal is full-wave-rectified to find its absolute value, then passed through a diode to a storage capacitor. When the absolute value of the voltage rises above the value stored on the capacitor, the diode will conduct and increase the stored voltage. When the voltage decreases, the capacitor will maintain the old value. Some method of discharging the capacitor is required so that a new peak value can be measured. In a true peak detector, this is accomplished by a solid-state switch. Practical peak detectors usually in-

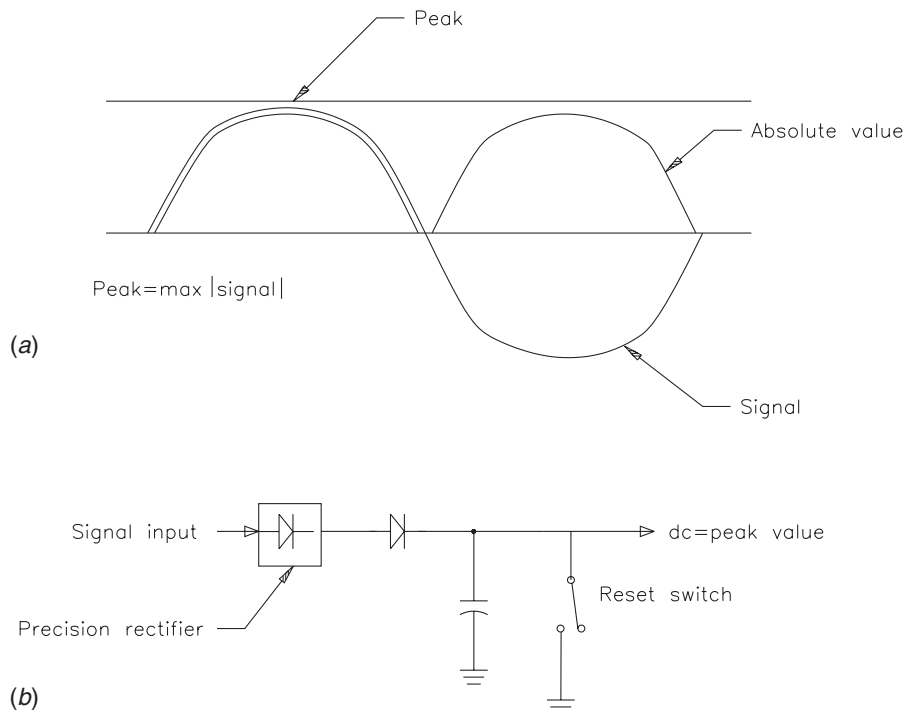


Figure 10.4 Peak voltage measurements: (a) illustration of peak detection, (b) peak measurement circuit.

clude a large-value resistor to discharge the capacitor gradually after the user has had a chance to read the meter.

The ratio of the true peak to the rms value is called the *crest factor*. For any signal but an ideal square wave, the crest factor will be greater than 1, as demonstrated in [Figure 10.5](#). As the measured signal becomes more peaked, the crest factor increases.

By introducing a controlled charge and discharge time, a *quasi-peak* detector is achieved. The charge and discharge times may be selected, for example, to simulate the transmission pattern for a digital carrier. The gain of a quasi-peak detector is normally calibrated so that it reads the same as an rms detector for sine waves.

The *peak-equivalent sine* is another method of specifying signal amplitude. This value is the rms level of a sine wave having the same peak-to-peak amplitude as the signal under consideration. This is the peak value of the waveform scaled by the correction factor 1.414, corresponding to the peak-to-rms ratio of a sine wave. Peak-equivalent sine is useful when specifying test levels of waveforms in distortion measurements. If the distortion of a device is measured as a function of amplitude, a point will be reached at which the output level cannot increase any further. At this point the peaks of the waveform will be clipped, and the distortion will rise rapidly with further increases in level. If another signal is used for distortion testing on the same device, it is desirable

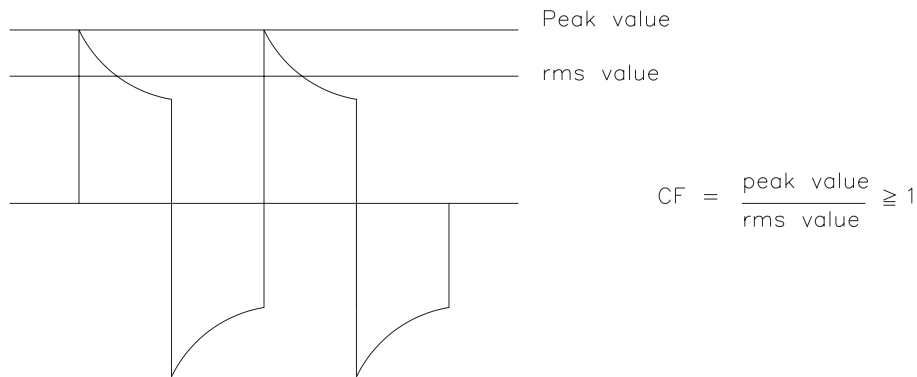


Figure 10.5 Illustration of the crest factor in voltage measurements.

that the levels at which clipping is reached correspond. Signal generators are normally calibrated in this way to allow changing between waveforms without clipping or readjusting levels.

Measurement Bandwidth

The bandwidth of the level-measuring instrument can have a significant effect on the accuracy of the reading. For a meter with a single-pole rolloff (one bandwidth-limiting component in the signal path), significant measurement errors can occur. Such a meter with a specified bandwidth of 1 MHz, for example, will register a 10 percent error in the measurement of signals at 500 kHz. To obtain 1 percent accurate measurements (disregarding other error sources in the meter), the signal frequency must be less than 100 kHz.

Figure 10.6 illustrates another problem associated with limited-bandwidth measuring devices. In the figure, a distorted sine wave is measured by two meters with different bandwidths. The meter with the narrower bandwidth does not respond to all the harmonics and gives a lower reading. The severity of this effect varies with the frequency being measured and the bandwidth of the meter; it can be especially severe in the measurement of wideband noise. Peak measurements are particularly sensitive to bandwidth effects. Systems with restricted low-frequency bandwidth will produce *tilt* in a square wave, and bumps in the high-frequency response will produce an *overshoot*. The effect of either will be an increase in the peak reading.

Meter Accuracy

Accuracy is a measure of how well an instrument quantifies a signal at a midband frequency. This sets a basic limit on the performance of the meter in establishing the absolute amplitude of a signal. It is also important to look at the *flatness* specification to

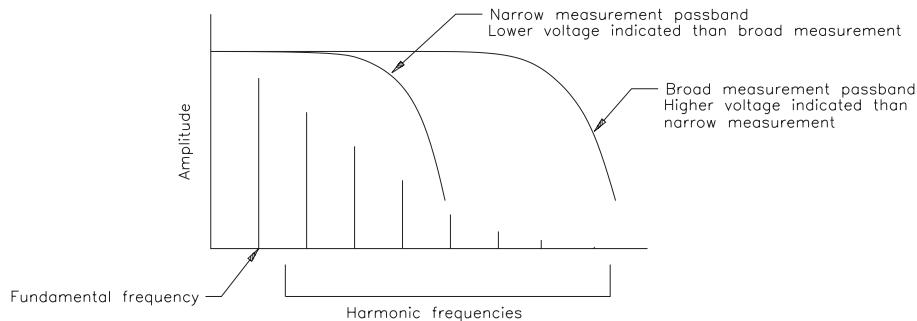


Figure 10.6 The effects of instrument bandwidth on voltage measurements.

see how well this performance is maintained with changes in frequency. Flatness describes how well the measurements at any other frequency track those at the reference. If a meter has an accuracy of 2 percent at 1 MHz and a flatness of 1 dB (10 percent) from 20 kHz to 20 MHz, the inaccuracy can be as great as 12 percent at 20 MHz.

Meters often have a specification of accuracy that changes with voltage range, being most accurate only in the range in which the instrument was calibrated. A meter with 1 percent accuracy on the 2 V range and 1 percent accuracy per step would be 3 percent accurate on the 200 V scale. Using the flatness specification given previously, the overall accuracy for a 100 V, 20 MHz sine wave is 14 percent. In many instruments, an additional accuracy derating is given for readings as a percentage of full scale, making readings at less than full scale less accurate.

However, the accuracy specification is not normally as important as the flatness. When performing frequency response or gain measurements, the results are relative and are not affected by the absolute voltage used. When measuring gain, however, the attenuator accuracy of the instrument is a direct error source. Similar comments apply to the accuracy and flatness specifications for signal generators. Most are specified in the same manner as voltmeters, with the inaccuracies adding in much the same manner.

RF Power Measurement

Measurement of RF power output typically is performed with a directional wattmeter, calibrated for use over a specified range of frequencies and power levels. Various grades of accuracy are available, depending upon the requirements of the application. The most accurate measurements usually are made by determining the temperature rise of the cooling through a calibrated dummy load. The power absorbed by the coolant, usually water, can be calculated from the following equation:

$$P = K \times Q \times \Delta T \quad (10.2)$$

Where:

Table 10.1 Common Decibel Values and Conversion Ratios

dB Value	Voltage Ratio	Power Ratio
0	1	1
+1	1.122	1.259
+2	1.259	1.586
+3	1.412	1.995
+6	1.995	3.981
+10	3.162	10
+20	10	100
+40	100	10,000
-1	0.891	0.794
-2	0.794	0.631
-3	0.707	0.501
-6	0.501	0.251
-10	0.3163	0.1
-20	0.1	0.01
-40	0.01	0.0001

A doubling of voltage will yield a value of 6.02 dB, and a doubling of power will yield 3.01 dB. This is true because doubling of voltage results in an increase in power by a factor of 4. [Table 10.1](#) lists the decibel values for some common voltage and power ratios.

RF engineers often express the decibel value of a signal relative to some standard reference instead of another signal. The reference for decibel measurements may be predefined as a power level, as in dBk (decibels above 1 kW), or it may be a voltage reference. Often, it is desirable to specify levels in terms of a reference transmission level somewhere in the system under test. These measurements are designated dB_r, where the reference point or level must be separately conveyed.

10.2.3 Noise Measurement

Noise measurements are specialized level measurements. Noise may be expressed as an absolute level by simply measuring the voltage at the desired point in the system. This approach, however, is often not very meaningful. Specifying the noise performance as the signal-to-noise ratio is a better approach. S/N is a decibel measurement of the noise level using the signal level measured at the same point as a reference. This makes measurements at different points in a system or in different systems directly comparable. A signal with a given S/N can be amplified with a perfect amplifier or attenuated with no change in the S/N. Any degradation in S/N at a later point in the system is the result of limitations of the equipment that follows.

Noise performance is an important parameter in the operation of any power vacuum tube amplifier or oscillator. All electric conductors contain free electrons that are in continuous random motion. It can be expected that, by pure chance, more electrons will be moving in one direction than in another at any instant. The result is that a voltage will be developed across the terminals of the conductor if it is an open circuit, or a current will be delivered to any connected circuit. Because this voltage (or current) varies in a random manner, it represents noise energy distributed throughout the frequency spectrum, from the lowest frequencies well into the microwave range. This effect is commonly referred to as *thermal-agitation noise*, because the motion of electrons results from thermal action. It is also referred to as *resistance noise*. The magnitude of the noise depends upon the following:

- The resistance across which the noise is developed
- Absolute temperature of the resistance
- Bandwidth of the system involved

Random noise, similar in character to that produced in a resistance, is generated in vacuum tubes as a result of irregularities in electron flow. Tube noise can be divided into the following general classes:

- *Shot effect*, representing random variations in the rate of electron emission from the cathode
- *Partition noise*, arising from chance variations in the division of current between two or more positive electrodes
- *Induced grid noise*, produced as a result of variations in the electron stream passing adjacent to a grid
- *Gas noise*, generated by random variations in the rate of ion production by collision
- *Secondary emission noise*, arising from random variations in the rate of production of secondary electrons
- *Flicker effect*, a low-frequency variation in emission that occurs with oxide-coated cathodes

Shot effect in the presence of space charge, partition noise, and induced grid noise is the principal source of tube noise that must be considered in RF work.

10.2.4 Phase Measurement

When a signal is applied to the input of a device, the output will appear at some later point in time. For sine wave excitation, this delay between input and output may be expressed as a proportion of the sine wave cycle, usually in degrees. One cycle is 360° , one half-cycle is 180° , and so on. This measurement is illustrated in [Figure 10.8](#). The

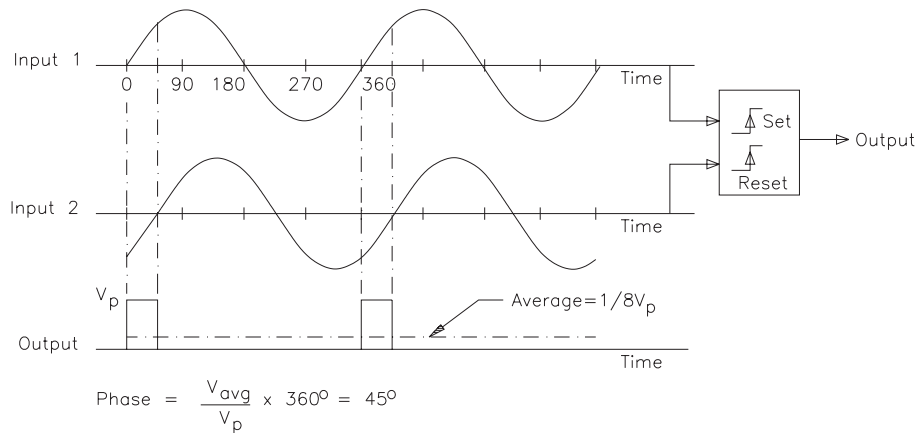


Figure 10.8 Illustration of the measurement of a phase shift between two signals.

phasemeter input signal number 2 is delayed from, or is said to be *lagging*, input number 1 by 45° .

Most RF test equipment checks phase directly by measuring the proportion of one signal cycle between zero crossings of the signals. Phase typically is measured and recorded as a function of frequency over a specified range. For most vacuum tube devices, phase and amplitude responses are closely coupled. Any change in amplitude that varies with frequency will produce a corresponding phase shift.

Relation to Frequency

When dealing with complex signals, the meaning of phase can become unclear. Viewing the signal as the sum of its components according to Fourier theory, a different value of phase shift is found at each frequency. With a different phase value on each component, the one to be used as the reference is unclear. If the signal is periodic and the waveshape is unchanged passing through the device under test, a phase value still may be defined. This may be done by using the shift of the zero crossings as a fraction of the waveform period. Indeed, most commercial phase-measuring instruments will display this value. However, in the case of differential phase shift with frequency, the waveshape will be changed. It is then impossible to define any phase-shift value, and phase must be expressed as a function of frequency.

Group delay is another useful expression of the phase characteristics of an RF device. Group delay is the slope of the phase response. It expresses the relative delay of the spectral components of a complex waveform. If the group delay is flat, all components will arrive at a given point together. A peak or rise in the group delay indicates that those components will arrive later by the amount of the peak or rise. Group delay is computed by taking the derivative of the phase response vs. frequency. Mathematically:

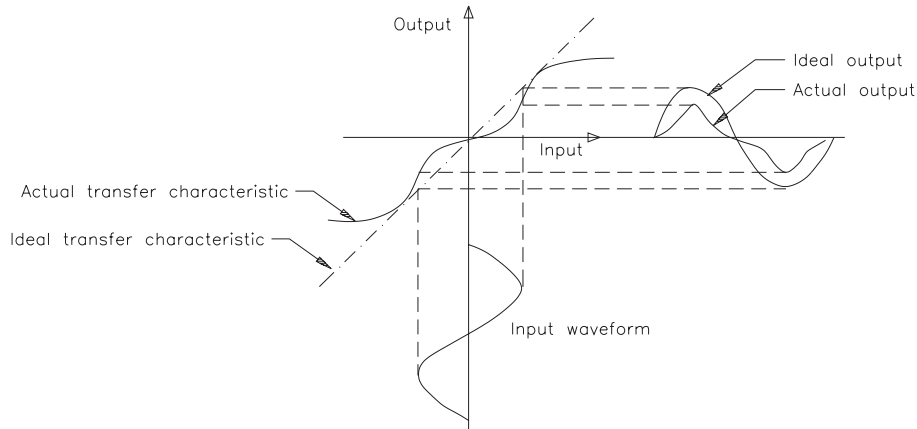


Figure 10.9 Illustration of total harmonic distortion (THD) measurement of an amplifier transfer characteristic.

$$\text{Group delay} = - \frac{(\text{phase at } f_2 - \text{phase at } f_1)}{f_2 - f_1} \quad (10.5)$$

This definition requires that phase be measured over a range of frequencies to give a curve that can be differentiated. It also requires that the phase measurements be performed at frequencies sufficiently close to provide a smooth and accurate derivative.

10.2.5 Nonlinear Distortion

Distortion is a measure of signal impurity, a deviation from ideal performance of a device, stage, or system. Distortion usually is expressed as a percentage or decibel ratio of the undesired components to the desired components of a signal. Distortion of a vacuum tube device or stage is measured by inputting one or more sine waves of various amplitudes and frequencies. In simplistic terms, any frequencies at the output that were not present at the input are distortion. However, strictly speaking, components caused by power line interference or another spurious signal are not distortion but, rather, noise. Many methods of measuring distortion are in common use, including harmonic distortion and several types of intermodulation distortion.

Harmonic Distortion

The transfer characteristic of a typical vacuum tube amplifier is shown in [Figure 10.9](#). The transfer characteristic represents the output voltage at any point in the signal waveform for a given input voltage; ideally this is a straight line. The output waveform is the projection of the input sine wave on the device transfer characteristic. A

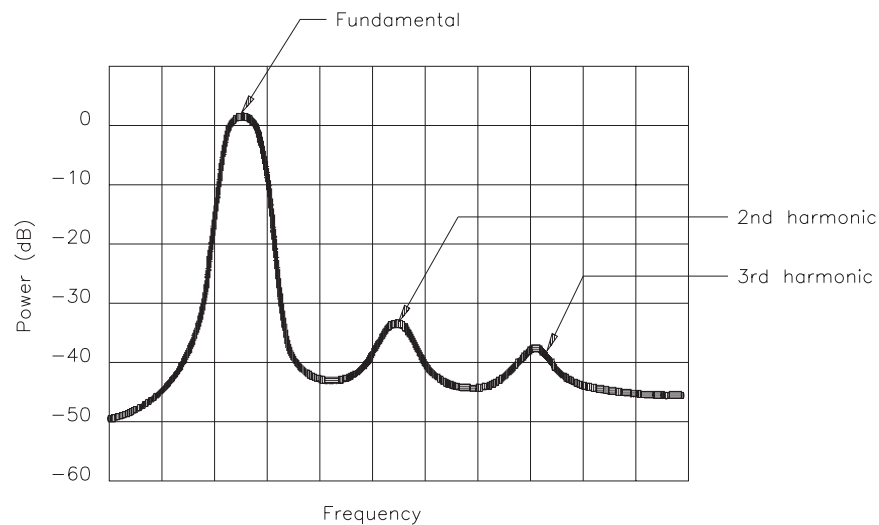


Figure 10.10 Example of reading THD from a spectrum analyzer.

change in the input produces a proportional change in the output. Because the actual transfer characteristic is nonlinear, a distorted version of the input waveshape appears at the output.

Harmonic distortion measurements excite the device under test with a sine wave and measure the spectrum of the output. Because of the nonlinearity of the transfer characteristic, the output is not sinusoidal. By using Fourier series, it can be shown that the output waveform consists of the original input sine wave plus sine waves at integer multiples (harmonics) of the input frequency. The spectrum of a distorted signal is shown in [Figure 10.10](#). The harmonic amplitudes are proportional to the amount of distortion in the device under test. The percentage of harmonic distortion is the rms sum of the harmonic amplitudes divided by the rms amplitude of the fundamental.

Harmonic distortion may be measured with a spectrum analyzer or a distortion test set. [Figure 10.11](#) shows the setup for a spectrum analyzer. As shown in [Figure 10.10](#), the fundamental amplitude is adjusted to the 0 dB mark on the display. The amplitudes of the harmonics are then read and converted to linear scale. The rms sum of these values is taken, which represents the THD.

A simpler approach to the measurement of harmonic distortion can be found in the notch-filter distortion analyzer, illustrated in [Figure 10.12](#). This device, commonly referred to as simply a distortion analyzer, removes the fundamental of the signal to be investigated and measures the remainder. [Figure 10.13](#) shows the notch-filter approach applied to a spectrum analyzer for distortion measurements.

The correct method of representing percentage distortion is to express the level of the harmonics as a fraction of the fundamental level. However, many commercial distortion analyzers use the total signal level as the reference voltage. For small amounts of distortion, these two quantities are essentially equivalent. At large values of distortion,

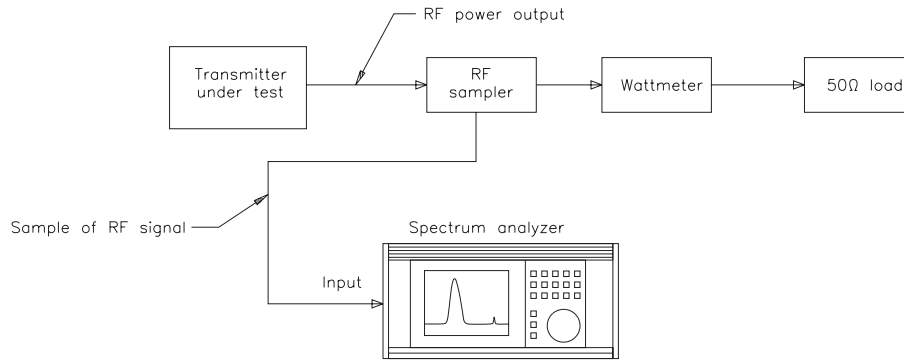


Figure 10.11 Common test setup to measure harmonic distortion with a spectrum analyzer.

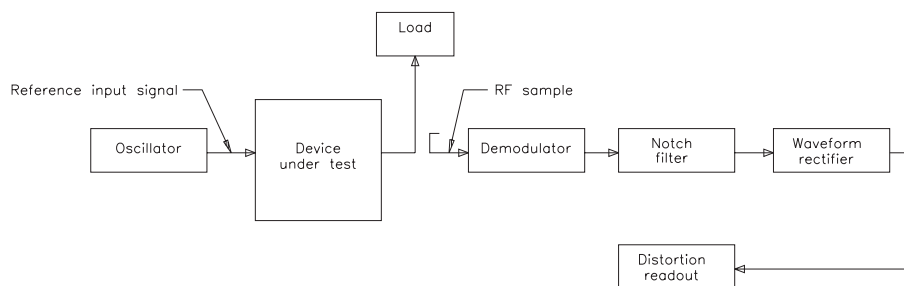


Figure 10.12 Simplified block diagram of a harmonic distortion analyzer.

however, the total signal level will be greater than the fundamental level. This makes distortion measurements on such units lower than the actual value. The relationship between the measured distortion and true distortion is given in [Figure 10.14](#). The errors are not significant until about 20 percent measured distortion.

Because of the notch-filter response, any signal other than the fundamental will influence the results, not just harmonics. Some of these interfering signals are illustrated in [Figure 10.15](#). Any practical signal contains some noise, and the distortion analyzer will include this noise in the reading. Because of these added components, the correct term for this measurement is *total harmonic distortion and noise* (THD+N). Additional filters are included on most distortion analyzers to reduce unwanted noise and permit a more accurate reading.

The use of a sine wave test signal and a notch-type distortion analyzer provides the distinct advantage of simplicity in both design and use. This simplicity has an addi-

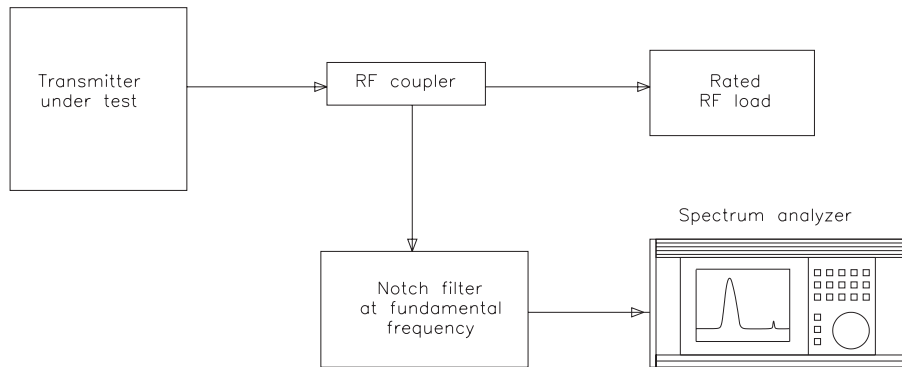


Figure 10.13 Typical test setup for measuring the harmonic and spurious output of a transmitter. The notch filter is used to remove the fundamental frequency to prevent overdriving the spectrum analyzer input and to aid in evaluation of distortion components.

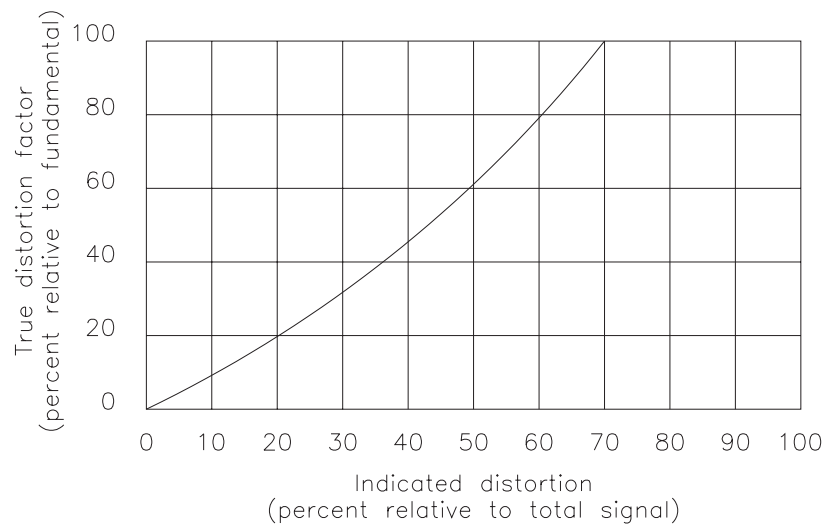


Figure 10.14 Conversion graph for indicated distortion and true distortion.

tional benefit in ease of interpretation. The shape of the output waveform from a notch-type analyzer indicates the slope of the nonlinearity. Displaying the residual components on the vertical axis of an oscilloscope and the input signal on the horizontal axis provides a plot of the deviation of the transfer characteristic from a best-fit straight line. This technique is diagrammed in [Figure 10.16](#). The trace will be a horizontal line for a perfectly linear device. If the transfer characteristic curves upward on positive input voltages, the trace will bend upward at the right-hand side.

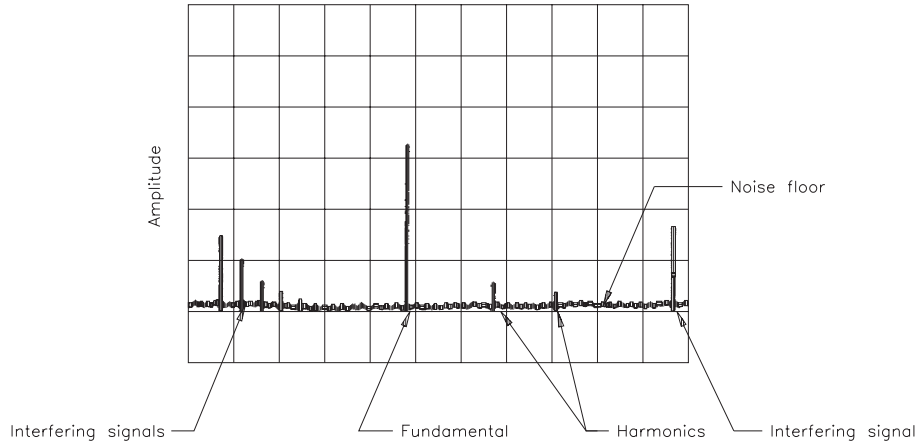


Figure 10.15 Example of interference sources in distortion and noise measurements.

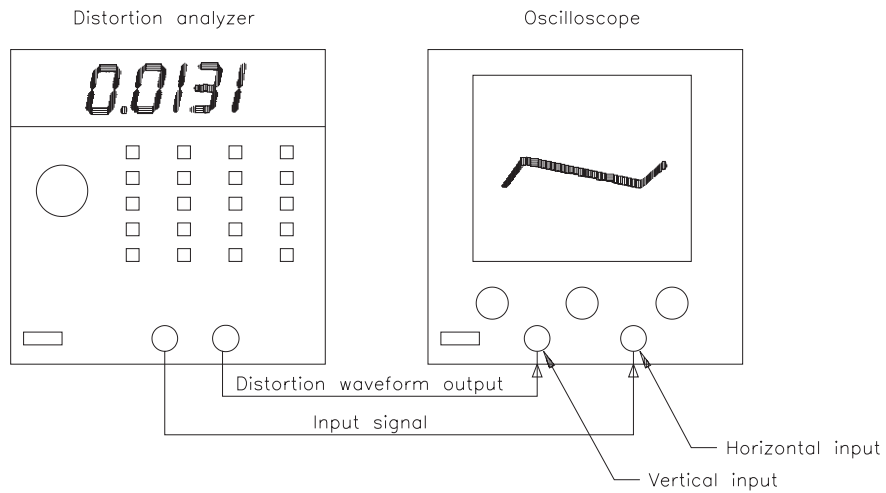


Figure 10.16 Transfer-function monitoring configuration using an oscilloscope and distortion analyzer.

Examination of the distortion components in real time on an oscilloscope allow observation of oscillation on the peaks of the signal and clipping. This is a valuable tool in the design and development of RF circuits, and one that no other distortion measurement method can fully match. Viewing the residual components in the frequency do-

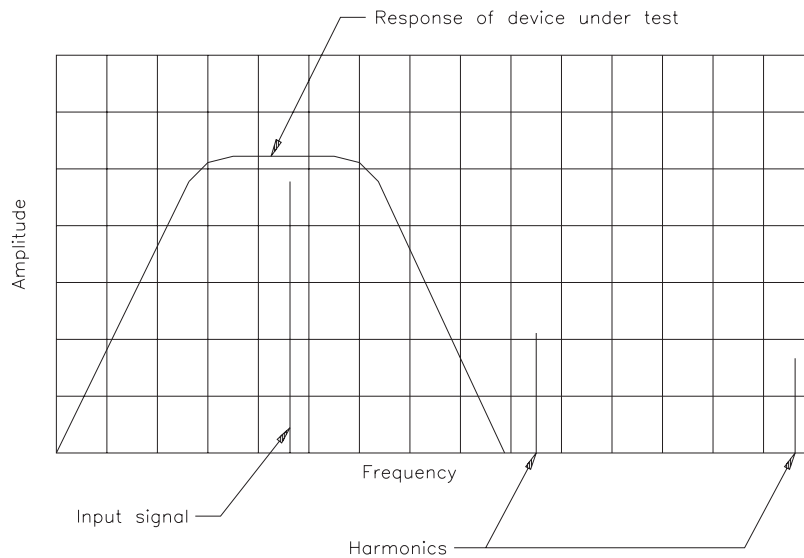


Figure 10.17 Illustration of problems that occur when measuring harmonic distortion in band-limited systems.

main by using a spectrum analyzer also reveals much information about the distortion mechanism inside the device or stage under test.

When measuring distortion at high frequencies, bandwidth limitations are an important consideration, as illustrated in Figure 10.17. Because the components being measured are harmonics of the input frequency, they may fall outside the passband of the device under test.

Intermodulation Distortion

Intermodulation distortion (IMD) describes the presence of unwanted signals that are caused by interactions between two or more desired signals. IMD, as it applies to an RF amplifier, refers to the generation of spurious products in a nonlinear amplifying stage. IMD can be measured using a distortion monitor or a spectrum analyzer.

Nonlinearities in a circuit can cause it to act like a mixer, generating the sum and difference frequencies of two signals that are present. These same imperfections generate harmonics of the signals, which then can be mixed with other fundamental or harmonic frequencies. Figure 10.18 shows the relationship of these signals in the frequency domain. The *order* of a particular product of IMD is defined as the number of “steps” between it and a single fundamental signal. Harmonics add steps. For example, the second harmonic = $2f_1 = f_1 + f_1$, which is two steps, making it a *second-order* product. The product resulting from $2f_1 - f_2$ is, then, a *third-order* product. Some of the IMD products resulting from two fundamental frequencies include:

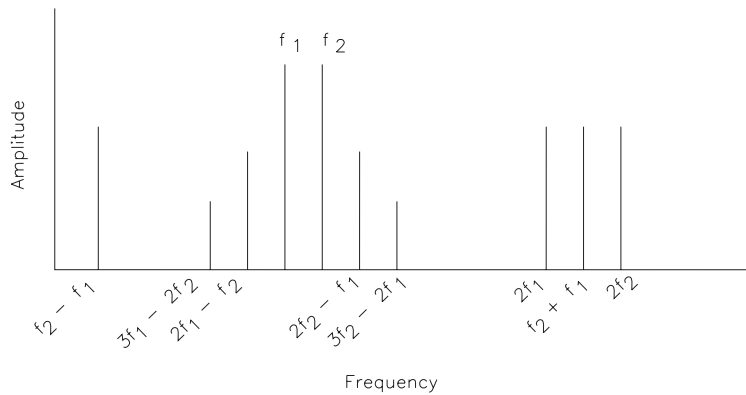


Figure 10.18 Frequency-domain relationships of the desired signals and low-order IMD products.

$$\text{second-order} = 2f_1 \quad (10.6)$$

$$= 2f_2 \quad (10.7)$$

$$= f_1 + f_2 \quad (10.8)$$

$$= f_1 - f_2 \quad (10.9)$$

$$\text{third-order} = 2f_1 + f_2 \quad (10.10)$$

$$= 2f_1 - f_2 \quad (10.11)$$

$$= 2f_2 + f_1 \quad (10.12)$$

$$= 2f_2 - f_1 \quad (10.13)$$

and so on.

The order is important because, in general, the amplitudes of IMD products fall off as the order increases. Therefore, low-order IMD products have the greatest potential to cause problems if they fall within the band of interest, or on frequencies used by other services or equipment. Odd-order IMD products are particularly troublesome because they fall closest to the signals that cause them, usually within the operating frequency band.

Measurement Techniques

A number of techniques have been devised to measure the intermodulation (IM) of two or more signals passing through a device simultaneously. The most common of

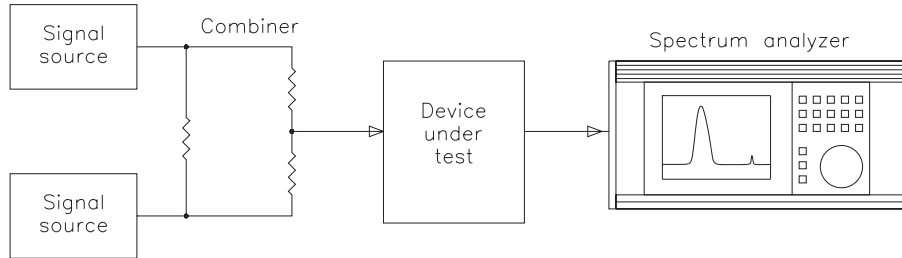


Figure 10.19 Test setup for measuring the intermodulation distortion of an RF amplifier using a spectrum analyzer.

these involves the application of a test waveform consisting of a low-frequency signal and a high-frequency signal mixed in a specified amplitude ratio. The signal is applied to the device under test, and the output waveform is examined for modulation of the upper frequency by the low-frequency signal. The amount by which the low-frequency signal modulates the high-frequency signal indicates the degree of nonlinearity. As with harmonic distortion measurement, this test may be done with a spectrum analyzer or a dedicated distortion analyzer. The test setup for a spectrum analyzer is shown in Figure 10.19. Two independent signal sources are connected using a power combiner to drive the device under test. The sources are set at the same output level, but at different frequencies. A typical spectrum analyzer display of the two-tone distortion test is shown in Figure 10.20. As shown, the third-order products that fall close to the original two tones are measured. This measurement is a common, and important, one because such products are difficult to remove by filtering.

The mechanisms of intermodulation as applied to interfering radiated signals are discussed later in this chapter.

Addition and Cancellation of Distortion Components

The addition and cancellation of distortion components in test equipment or the device under test is an often-overlooked problem in measurement. Consider the examples in Figure 10.21 and 10.22. Assume that one device under test has a transfer characteristic similar to that diagrammed at the top left of Figure 10.21a, and another has the characteristic diagrammed at the center. If the devices are cascaded, the resulting transfer-characteristic nonlinearity will be magnified as shown. The effect on sine waves in the time domain is illustrated in Figure 10.21b. The distortion components generated by each nonlinearity are in phase and will sum to a component of twice the original magnitude. However, if the second device under test has a complementary transfer characteristic, as shown in Figure 10.22, quite a different result is obtained. When the devices are cascaded, the effects of the two curves cancel, yielding a straight line for the transfer characteristic (Figure 10.22a). The corresponding distort-

Range: -10 dBm
 Res BW: 200 Hz
 Marker: 10,010,500 Hz -15.07 dBm
 Sweep time 1.23 s

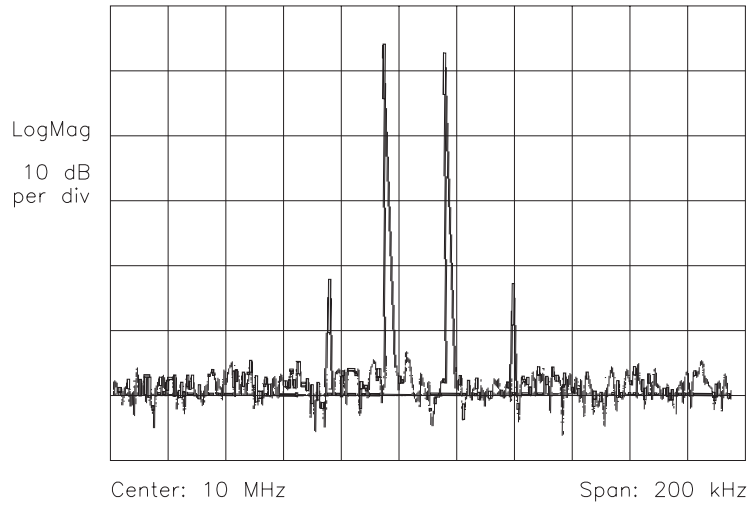


Figure 10.20 Typical two-tone IMD measurement, which evaluates the third-order products within the bandpass of the original tones.

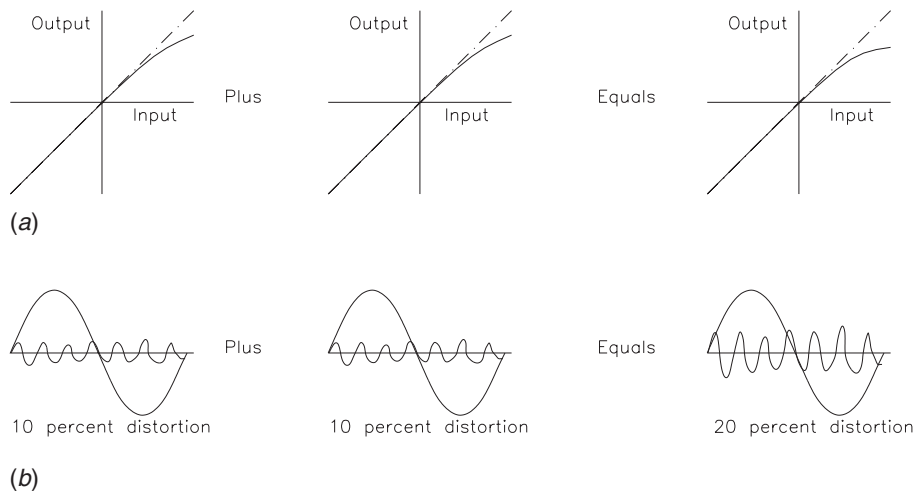


Figure 10.21 Illustration of the addition of distortion components: (a) addition of transfer-function nonlinearities, (b) addition of distortion components.

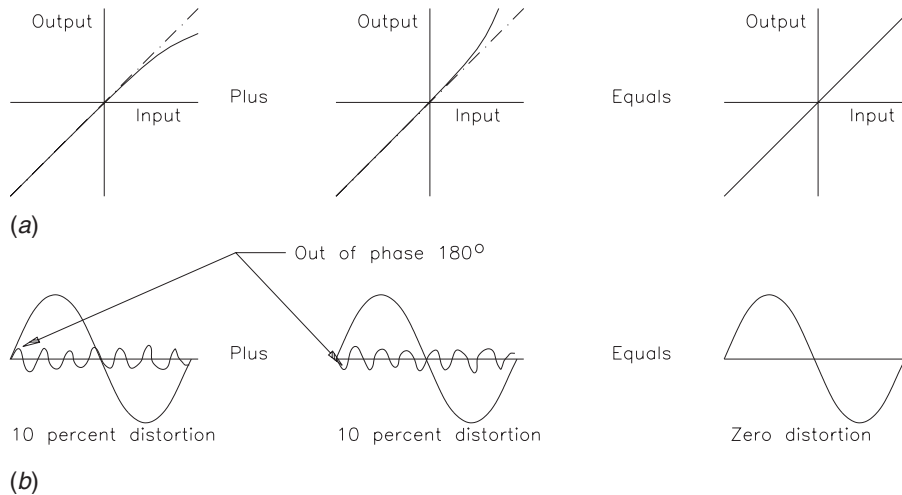


Figure 10.22 Illustration of the cancellation of distortion components: (a) cancellation of transfer-characteristic nonlinearities, (b) cancellation of the distortion waveform.

tion products are out of phase, resulting in no measured distortion components in the final output (Figure 10.22b).

This problem is common at low levels of distortion, especially between the test equipment and the device under test. For example, if the test equipment has a residual of 0.002 percent when connected to itself, and readings of 0.001 percent are obtained from the circuit under test, cancellations are occurring. It is also possible for cancellations to occur in the test equipment itself, with the combined analyzer and signal generator giving readings lower than the sum of their individual residuals. If the distortion is the result of even-order (asymmetrical) nonlinearity, reversing the phase of the signal between the offending devices will change a cancellation to an addition. If the distortion is the result of an odd-order (symmetrical) nonlinearity, phase inversions will not affect the cancellation.

Intermodulation Precorrection Techniques

In a practical system, amplifiers cannot be absolutely linear; all linear amplifiers exhibit intermodulation distortion to some extent, relative to the amplitudes of the parent signals [1]. Adequate linearity is judged by the application, but in general, the in-band odd-order products should be as low as the state of the art permits because they are in-band and thus cannot be filtered out.

All intermodulation products are the result of undesirable time domain multiplication of the matrix of input signals to a linear amplifier. All harmonic components are the result of the same process, only here, an individual signal is multiplied by itself. For

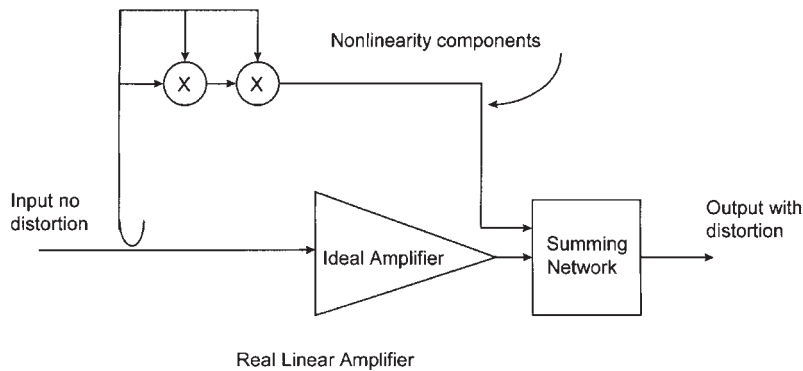


Figure 10.23 Basic model of an intermodulation correction circuit. (After [1].)

second order intermodulation, the ideal linear amplifier is modified by modeling a single mixer feeding a time domain multiplied signal around the amplifier. For third order intermodulation, the ideal linear amplifier is modified by modeling two mixers feeding a time domain multiplied signal around the amplifier, as shown in [Figure 10.23](#).

Because conventional amplitude modulation contains at least three parent signals (the carrier plus two sidebands), AM amplification has always been fraught with in-band intermodulation distortion. A common technique to correct for these distortion products is to precorrect the power amplifier using the following process:

- Intentionally generate an anti-intermodulation signal by mixing the three input signals together
- Adjust the amplitude of the anti-IM product to match that of the real IM product
- Invert the phase
- Delay the correction signal or the parent signals in time so as to permit proper alignment
- Add the correction signal into the main parent signal path to cause cancellation in the output circuit of the amplifier

This is a well-known technique used in countless applications.

Each intermodulation product may have its own anti-intermodulation product intentionally generated to cause targeted cancellation. Even if one or more parent signals contains modulation of some type—AM, FM, or PM—the legitimate modulation signal is often included in the intended mixing circuit so that even the sidebands produced around an intermodulation distortion product are canceled.

The precorrection circuit typically consists of mixer devices with a local oscillator (LO) port, an RF port, and an IF port. The LO port is fed with the parent signal, such as the carrier, which contains no modulation. Its purpose is to hard-switch the mixer di-

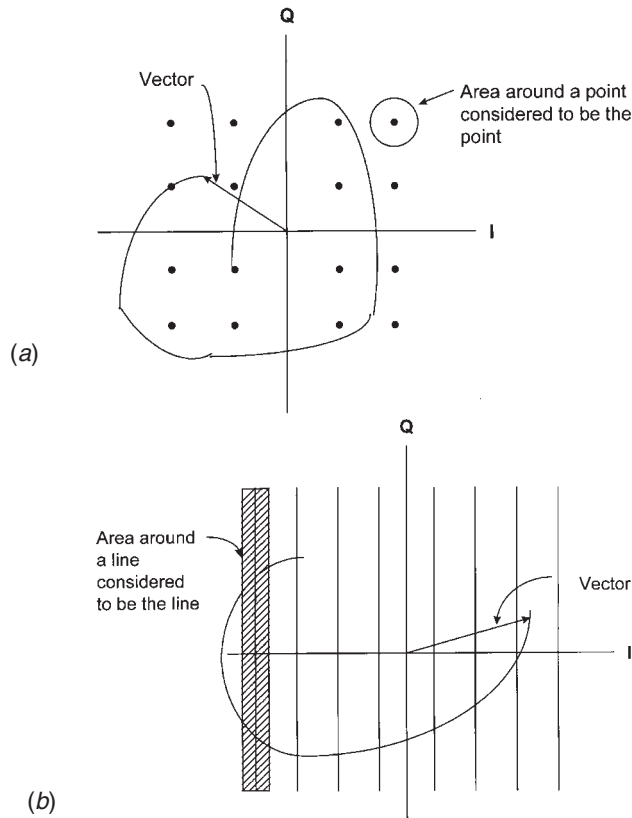


Figure 10.24 Digital decoding process: (a) 16-QAM, (b) 8-VSB. (After [1].)

odes, causing frequency translation of the signal fed to the RF port to the sum and difference frequencies of the LO and RF input signals exiting at the IF port. The intermodulation product to be canceled is either a result of the sum or the difference signal so that the product of no concern is filtered out. The resulting intended mixer product is fed to the IF port of a second mixer where LO injection is from the third parent signal. The third order intended mixing product is then available at the second mixer RF port, and will be at the frequency of the offending IM and coherent with it.

Digital signals suitable for radio transmission actually remain analog, but are termed *digital* because of the nature of the information that they carry. A modern digital radio transmission using quadrature amplitude modulation (QAM) or n -level vestigial sideband modulation (e.g., 8-VSB) consists of a vector signal constantly moving in phase angle and magnitude relative to some reference. The vector signal is smooth and without discontinuity so that it is well contained within a finite bandwidth. Its digital nature stems from decoding the proper point or spatial area around a point (QAM) of a

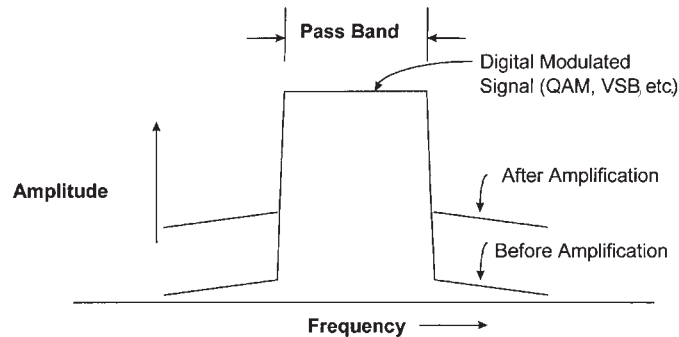


Figure 10.25 Digitally modulated RF signal spectrum before and after amplification. (After [1].)

finite number of legitimate points, or the proper line (VSB) or spatial area around a line of a finite number of legitimate lines, that the vector is passing through. Figure 10.24 illustrates this process. In either case, the area is assigned a discrete digital value, which represents a portion of the information being transmitted. Spatial location is generally expressed in terms of rectangular coordinates with I (in phase) and Q (quadrature phase) components.

Because the vector appears to move at random from point to point, especially resulting from the use of an intentional randomizer circuit, the frequency distribution within the passband of the signal is constant, flat, and favors no particular frequency. Figure 10.25 illustrates the frequency distribution with the sidebands shown before and after amplification. As a result, the transmitted signal is said to be “noise-like.” As such, it takes on the nature of noise and need not be periodic.

For this case, there is no identifiable set of parent components upon which to obtain signals for the purpose of mixing to generate an anti-intermodulation signal for cancellation purposes. There is no carrier signal to feed into the LO port of a mixer for hard switching to generate sum and difference frequencies. Because the signal is noise-like, there is no correlation of any piece of passband spectrum with any other. Generating a precorrection signal, thus, is not possible using the process described for the discrete signal analog case.

10.3 Vacuum Tube Operating Parameters

Direct current input power to the final RF output stage is a function of the applied plate or collector voltage and the plate or collector current of the final tube(s) [2]. Test points or built-in meters usually are provided for checking the dc voltage and current levels of the final amplifier.

Transmitter deviation or modulation must be set properly. A deviation level set too high or too low adversely affects radio range. If the deviation is set too low, the sig-

nal-to-noise ratio of the overall system is diminished. If the deviation is set too high, the transmitter signal may exceed the receiver's modulation acceptance bandwidth, causing severe distortion on modulation peaks and out-of-band spurious radiation.

The difference between the positive and negative excursions of the deviated waveform is defined as *modulation dissymmetry*. Theoretically, positive and negative excursions should be perfectly symmetrical. In practice, however, there is usually some dissymmetry. In a multichannel phase-modulated transmitter, modulation dissymmetry may be worse on some channels than others as a result of modulator tuning. If the modulator tuning is peaked on one channel, for example, response on another channel may be compromised.

10.3.1 Stage Tuning

A power output stage can be tuned in any number of ways [2]. Determining factors include the type of device, circuit design, and operational objectives. A PA can be tuned to optimize bandwidth, linearity, efficiency, and other parameters. It is often necessary to compromise one parameter to enhance another. This tradeoff is dictated by the application.

Power Grid Tubes

When optimizing a tetrode or pentode RF amplifier for proper excitation and loading, the general tuning procedure varies depending upon whether the screen voltage is taken from a fixed supply or a dropping resistor supply with poor regulation [2]. If both the screen supply and grid bias are from fixed sources with good regulation, the plate current is almost entirely controlled by RF excitation. The loading is then varied until the maximum power output is obtained. Following these adjustments, the excitation is trimmed, along with the loading, until the desired control and screen grid currents are obtained.

In the case of an RF amplifier where both the screen and grid bias are taken from sources with poor regulation, the stage will tune in a manner similar to a triode RF power amplifier. The plate current will be controlled principally by varying the loading, and the excitation trimmed to yield the desired control grid current. In this case, the screen current will be almost entirely set by the choice of the dropping resistor. Excitation and loading will vary the screen voltage considerably, and these should be trimmed to yield a desired screen voltage.

The grounded-grid amplifier has been used for many years, and, with the advent of high-power zero-bias triodes, this configuration has become more common. Adjustment of the excitation and loading of a grounded-grid RF amplifier requires a slightly different procedure. The plate voltage (plate and screen voltage in the case of a tetrode or pentode) must be applied before the excitation. If this precaution is not followed, the control grid may sustain damage. The loading is increased as the excitation is increased. When the desired plate current is reached, note the power output. The loading can be reduced slightly and the excitation increased until the plate current is the same as before. If the power output is less than before, a check can be made with increased load-

ing and less excitation. By proper trimming, the proper grid current, plate current, and optimum power output can be obtained. In a grounded-grid circuit, the cathode or input circuit is in series with the plate circuit. Because of this arrangement, any change made in the plate circuit will affect the input circuit. Therefore, the driver amplifier does not see its designed load until the driven stage is up to full plate current.

Typical operating values provide a starting point for tuning efforts. As long as the maximum limits are not exceeded, a wide range of tuning options is available.

Screen Voltage

Typical operating values given on the data sheet for tetrode and pentode devices include recommended screen voltages [2]. These values are not critical for most applications; they are chosen as a convenient value consistent with low driving power and reasonable screen dissipation. If lower values of screen voltage are used, more driving voltage will be required on the grid to obtain the same plate current. If higher values of screen voltage are used, less driving voltage will be required. Thus, high power gain can be achieved, provided the circuit exhibits adequate stability. Be careful not to exceed the screen dissipation limit. The value of screen voltage can be chosen to suit available power supplies or amplifier conditions.

The published characteristic curves of tetrodes and pentodes are shown for commonly used screen voltages. Occasionally, it is desirable to operate the tetrode or pentode at some screen voltage other than that shown on the characteristic curves. It is a relatively simple matter to convert the published curves to corresponding curves at a different screen voltage. This conversion is based on the fact that if all interelectrode voltages are either raised or lowered by the same relative amount, the shape of the voltage field pattern is not altered, and neither is the current distribution. The current lines simply take on new proportionate values in accordance with the *three-halves power law*. This method fails only where insufficient cathode emission or high secondary emission affects the current values.

For example, if the characteristic curves are shown at a screen voltage of 250 V and it is desirable to determine conditions at 500 V on the screen, all voltage scales should be multiplied by the same factor that is applied to the screen voltage (2 in this case). The 1 kV plate voltage line thus becomes 2 kV, the 50 V grid voltage line becomes 100 V, and so on. All current lines then assume new values in accordance with the $3/2$ power law. Because the voltage was increased by a factor of 2, the current lines all will be increased in value by a factor of $2^{3/2}$ or 2.8. All current values are then multiplied by the factor of 2.8. The 100 mA line thus becomes a 280 mA line, and so on.

Likewise, if the screen voltage given on the characteristic curve is higher than the conditions desired, the voltage should be reduced by the same factor that is used to obtain the desired screen voltage. Correspondingly, the current values all will be reduced by an amount equal to the $3/2$ power of this factor. Commonly used factors are given in [Table 10.2](#).

Table 10.2 Three-Halves Current Values of Commonly Used Factors

Voltage Factor	Current Factor
0.25	0.125
0.50	0.35
0.75	0.65
1.00	1.00
1.25	1.40
1.50	1.84
1.75	2.30
2.00	2.80
2.25	3.40
2.50	4.00
2.75	4.60
3.00	5.20

Back-Heating by Electrons

Back-heating of the cathode is an undesirable action involving the motion of electrons within the tube at VHF and UHF [2]. Because the time of flight of electrons from the cathode through the grid structure to the plate becomes an appreciable part of the cycle at high frequencies, electrons can be stopped in flight and turned back by the rapidly changing grid voltage. Under these conditions, the electrons are deflected from their normal paths and given excess energy with which they bombard the cathode and other portions of the tube structure. This effect can be aggravated by the choice of operating conditions to the extent that destructive effects occur. The tube can be destroyed within a few minutes under severe conditions.

Fortunately, the conditions that tend to minimize this back-bombardment by electrons also characterize minimum driving voltage. The tendency of electrons to turn back in flight is reduced by using the lowest possible RF grid voltage on the tube. This is achieved by using the lowest possible dc grid bias. In tetrodes, this effect is inherently lower because of the action of the screen in accelerating the electrons toward the anode. A tetrode also permits the use of smaller grid voltages. Consequently, under favorable conditions, the number of electrons turned back to heat the cathode and tube structure can be kept to a practical, low level. In addition to low dc grid bias, high screen voltage is desirable.

The plate circuit should always operate with heavy loading (low external plate impedance) so that the minimum instantaneous value of plate voltage will stay sufficiently positive to continue accelerating electrons to the anode. For this reason, the longest tube life is realized when the tetrode amplifier is heavily loaded, as indicated by having small values of dc screen and dc control grid current.

Never operate a tube with light plate loading. If the plate load is removed so that the minimum instantaneous plate voltage tends to fall to values approximating cathode potential (as it must do when the loading is removed completely and excitation is present), the number of electrons turned back can be destructive to the tube. It has been found that under conditions of no loading, the electron bombardment and increased electric field heating of the insulating portion of the tube is often sufficient to cause suck-in of the glass or cracking of the ceramic envelope. Automatic protection should be provided to remove all voltages from the tube if the plate circuit loading becomes too light for the amount of excitation applied.

Note that parasitic oscillations seldom are loaded heavily, as indicated by the high grid currents often observed during such self-oscillation. Thus, excessive RF plate voltages are developed that, at VHF, can be as damaging as unloaded operation on a VHF fundamental frequency. If unloaded VHF parasitic oscillations are present simultaneously with apparently satisfactory operation on the fundamental, unexplained reduction of tube life may result.

Occasionally, an output line circuit may resonate simultaneously to a harmonic frequency as well as to the fundamental frequency. The higher resonant modes of practical line circuits are not normally harmonically related. Sometimes, however, the tuning curve of a mode will cross the fundamental tuning curve and, at that point, the circuit will build up resonant voltages at both the harmonic and the fundamental frequency. The harmonic resonance usually is lightly loaded and the damaging action is similar to that of lightly loaded parasitic or fundamental. Again, while operation of the tube and circuit on the fundamental may appear normal, but with lower than expected efficiency, damage can occur to the tube.

In addition to operating the tube with minimum bias, high screen voltage, and heavy loading on the plate circuit, some degree of compensation for the remaining back-heating of the cathode may be required. This can be accomplished by lowering the filament voltage or heater voltage until the cathode operates at a normal temperature. In the case of tetrodes and pentodes, by taking precautions necessary to minimize back-bombardment by electrons, the required compensation for back-heating of the cathode is not significant and may often be neglected.

10.3.2 Amplifier Balance

In a push-pull RF amplifier, lack of balance in the plate circuit or unequal plate dissipation usually is caused by a lack of symmetry in the RF circuit [2]. Normally, the tubes are similar enough that such imbalance is not associated with the tubes and their characteristics. This can readily be checked by interchanging the tubes in the sockets (provided both tubes have common dc voltages to plate, screen, and grid) and observing whether the unbalanced condition remains at the socket location or moves with the tube. If it remains at the socket location, the circuit requires adjustment. If appreciable imbalance is associated with the tube, it is likely that one tube is abnormal and should be replaced.

The basic indicators of balance are the plate current and plate dissipation of each tube. It is assumed that the circuit applies the same dc plate voltage, dc screen voltage

(in the case of a tetrode or pentode), and dc grid bias to each tube from common supplies. Also, it is initially assumed that the plate circuit is mechanically and electrically symmetrical (or approximately so).

Imbalance in a push-pull RF amplifier is usually the result of unequal RF voltages applied to the grids of the tubes, or the RF plate circuit applying unequal RF voltages to the plates of the tubes. Correction of this situation involves first balancing grid excitation until equal dc plate currents flow in each tube. Next, the RF plate circuit is adjusted until equal plate dissipation appears on each tube, or equal RF plate voltage is observed. The balance of plate current is a more important criterion than equality of screen current (in a tetrode or pentode) or grid current. This results from the fact that tubes tend to be more uniform in plate current characteristics. However, the screen current also is sensitive to a lack of voltage balance in the RF plate circuit and may be used as an indicator. If the tubes differ somewhat in screen current characteristics, and the circuit has common dc supply voltages, final trimming of the plate circuit balance may be made by interchanging tubes and adjusting the circuit to give the same screen current for each tube, regardless of its location.

Note that the dc grid current has not been used as an indicator of balance of the RF power amplifier. It is probable that after following the foregoing procedure, the grid currents will be fairly well balanced, but this condition, in itself, is not a safe indicator of grid excitation balance.

10.3.3 Parallel Tube Amplifiers

The previous discussion has been oriented toward the RF push-pull amplifier. The same comments can be directed to parallel tube RF amplifiers [2]. Balance—the condition in which each tube carries its fair share of the load—still must be considered. In low-frequency amplifiers operating in class AB₁ or class AB₂, the idle dc plate current per tube should be balanced by separate bias adjustments for each tube. In many cases, some lack of balance of the plate currents will have negligible effect on the overall performance of the amplifier.

When tubes are operating in the idle position close to cutoff, an exact balance is difficult to achieve because plate current cannot be held to a close percentage tolerance. At this point, the action of the plate and screen voltages is in delicate balance with the opposing negative grid voltage. The state of balance is indicated by the plate current. Minor variations of individual grid wires or variations in the diameter of grid wires can upset the balance. It is practically impossible to control such minor variations in manufacture.

10.3.4 Harmonic Energy

A pulse of plate current delivered by the tube to the output circuit contains components of the fundamental and most harmonic frequencies [2]. To generate output power that is an harmonic of the exciting voltage applied to the control grid, it is merely necessary to resonate the plate circuit to the desired harmonic frequency. To optimize the performance of the amplifier, it is necessary to adjust the angle of plate

current flow to maximize the desired harmonic. The shorter the length of the current pulse, in the case of a particular harmonic, the higher the plate efficiency. However, the bias, exciting voltage, and driving power must be increased. Also, if the pulse is too long or too short, the output power will decrease appreciably.

The plate circuit efficiency of tetrode and pentode harmonic amplifiers is quite high. In triode amplifiers, if feedback of the output harmonic occurs, the phase of the voltage feedback usually reduces the harmonic content of the plate pulse, and thereby lowers the plate circuit efficiency. Because tetrodes and pentodes have negligible feedback, the efficiency of a harmonic amplifier usually is comparable to that of other amplifier types. Also, the high amplification factor of a tetrode or pentode causes the plate voltage to have little effect on the flow of plate current. A well-designed tetrode or pentode also permits large RF voltages to be developed in the plate circuit while still passing high peaks of plate current in the RF pulse. These two factors help to further increase plate efficiency.

Controlling Harmonics

The previous discussion of harmonics has been for the case where harmonic power in the load is the objective. The generation and radiation of harmonic energy must be kept to a minimum in a *fundamental frequency* RF amplifier [2]. The ability of the tetrode and pentode to isolate the output circuit from the input circuit over a wide range of frequencies is important in avoiding feedthrough of harmonic voltages from the grid. An important part of this shielding is the result of the basic physical design of the devices. The following steps permit reduction of unwanted harmonic energy in the output circuit:

- Minimize the circuit impedance between plate and cathode. This objective may be achieved by having some or all of the tuning capacitance of the resonant circuit close to the tube.
- Completely shield the input and output compartments of the stage.
- Use inductive output coupling from the resonant plate circuit, and possibly a capacitive or Faraday shield between the coupling coil and the tank coil. A high-frequency attenuating circuit, such as a *Pi* or *Pi-L* network, also may be used.
- Incorporate low-pass filters on all supply leads and wires leading to the output and input compartments.
- Use resonant traps for particular frequencies as required.
- Install a low-pass filter in series with the output transmission line.

10.3.5 Klystron Tuning Considerations

The output power stability of a klystron is sensitive to changes in the principal operating parameters, including:

- Beam input power

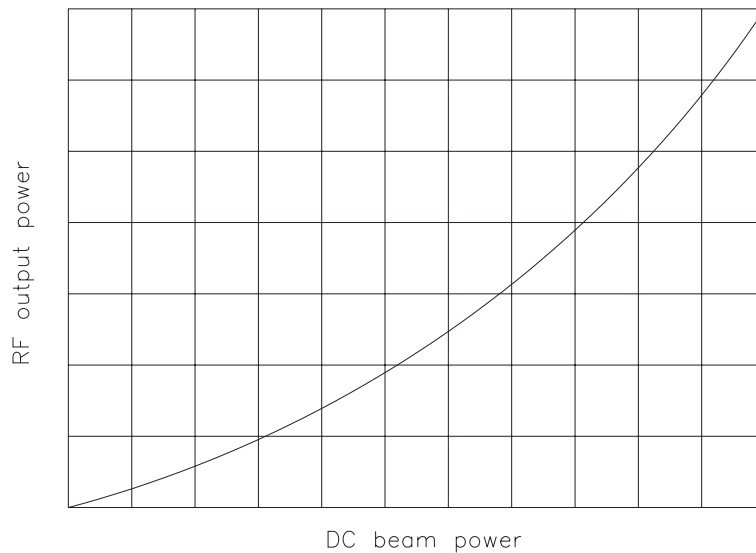


Figure 10.26 RF output power variation as a function of dc beam input power for a klystron.

- RF drive power
- RF drive frequency
- Device tuning
- Magnetic fields in the vicinity of the tube
- VSWR of the load

Varying any of these parameters will affect the RF output power response.

Figure 10.26 illustrates how changes in the dc beam input power affect the RF output power of a klystron under constant RF drive conditions. Small changes in dc input power produce marked changes in RF output power. Figure 10.27 shows RF output power as a function of RF drive power applied to the tube. From this curve, it can be seen that when the RF drive power is low, the RF output power is low. As the level of RF drive power increases, RF output power increases until an optimum operating point is reached. Beyond this point, further increases in RF drive power result in reduced RF output power.

Klystrons are said to be *saturated* at the point where a further increase in RF drive only decreases the RF output power. To obtain maximum RF output from a klystron, sufficient RF drive must be applied to the tube to reach the point of saturation. Operating at RF drive levels beyond saturation will only overdrive the klystron, decrease RF output power, and increase the amount of beam interception (body current) at the drift tube.

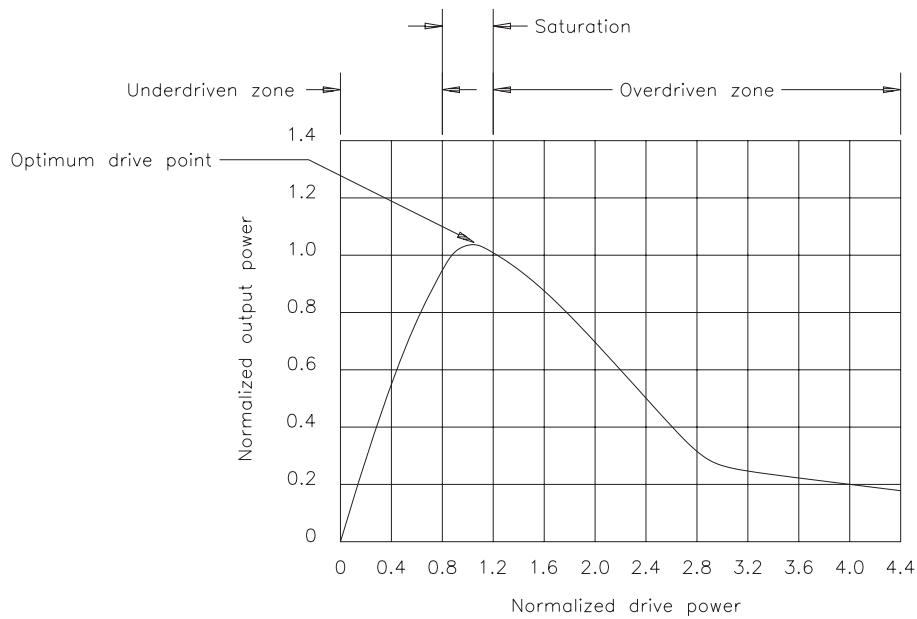


Figure 10.27 RF output power of a klystron as a function of RF drive power.

Tuning of any cavity of a multicavity klystron will affect the overall shape of the amplitude-response curve. There are two common methods of tuning klystrons:

- High-efficiency narrowband tuning
- Broadband tuning, yielding lower gain and efficiency

Figure 10.28 shows how RF output power changes with various levels of RF drive applied to a klystron under different tuning conditions. The RF power at point *A* represents the drive saturation point for a synchronously tuned tube. Point *B* shows a new point of saturation that is reached by tuning the *penultimate* (next to the last) cavity to a somewhat higher frequency. By tuning the penultimate cavity still further, point *C* is reached. At point *D*, increasing the penultimate cavity frequency no longer increases RF output power; instead, it reduces the power delivered by the tube (curve *E*).

10.3.6 Intermodulation Distortion

Intermodulation distortion products in linear power amplifier circuits can be caused by either amplitude gain nonlinearity or phase shift with change in input signal level. Intermodulation distortion products appear when the RF signal has a varying envelope amplitude. A single continuous-frequency wave will be amplified by a fixed amount and shifted in phase by a fixed amount. The nonlinearity of the amplifier will

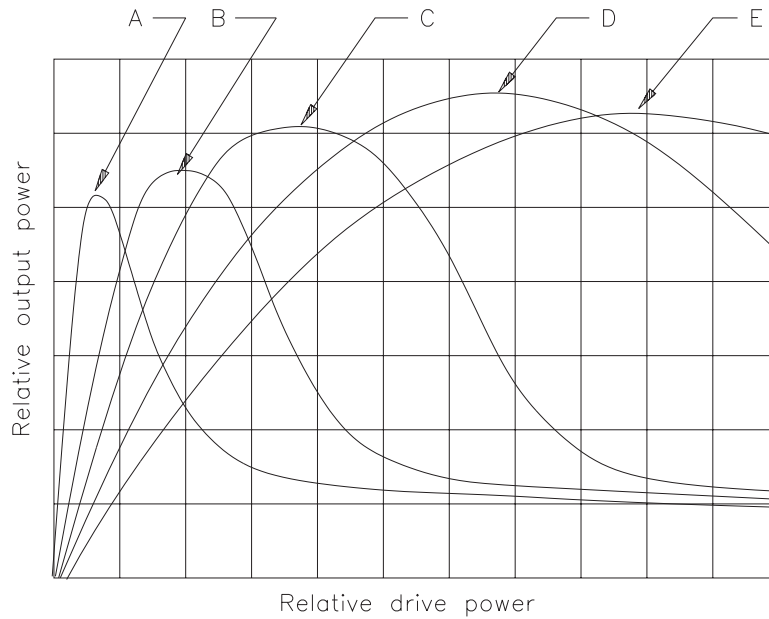


Figure 10.28 Output power variation of a klystron as a function of drive power under different tuning conditions.

produce only harmonics of the input wave. As the input RF wave changes, the nonlinearity of the amplifier will cause undesirable intermodulation distortion.

When an RF signal with varying amplitude is passed through a nonlinear device, many new products are produced. The frequency and amplitude of each component can be determined mathematically. The nonlinear device can be represented by a power series expanded about the zero-signal operating point. Intermodulation distortion energy is wasted and serves no purpose other than to cause interference to adjacent channels. Intermodulation distortion in a power amplifier tube is principally the result of its transfer characteristics. An ideal transfer characteristic curve is shown in [Figure 10.29](#).

Even-order products do not contribute to the intermodulation distortion problem because they usually fall outside the amplifier passband. Therefore, if the transfer characteristic produces an even-order curvature at the small-signal end of the curve (from point *A* to point *B*) and the remaining portion of the curve (point *B* to point *C*) is linear, the tube is considered to have an ideal transfer characteristic. If the operating point of the amplifier is set at point 0 (located midway horizontally between point *A* and point *B*), there will be no distortion in a class AB amplifier. However, no tube has this idealized transfer characteristic. It is possible, by manipulation of the electron ballistics within a given tube structure, to alter the transfer characteristic and minimize the distortion products.

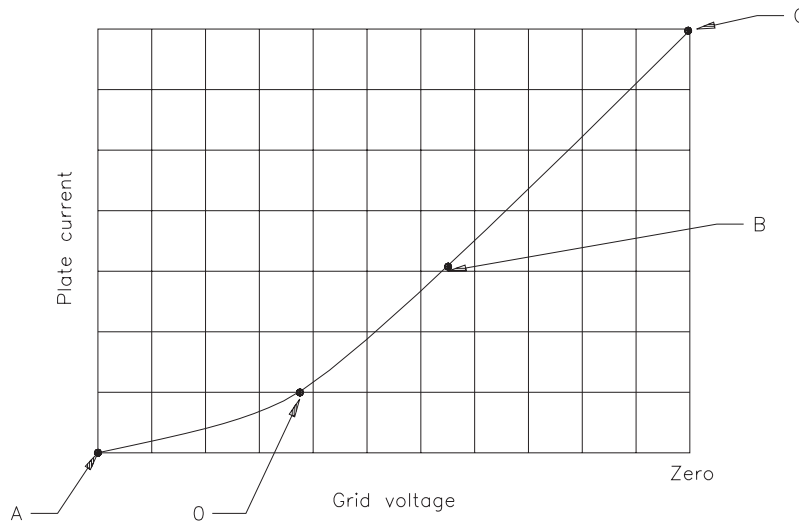


Figure 10.29 Ideal grid-plate transfer curve for class AB operation.

Generation of RF Intermodulation Products

When two or more transmitters are coupled to each other, new spectral components are produced by mixing of the fundamental and harmonic terms of each of the desired output frequencies.¹ The most common form of coupling results from transmitting antennas that are physically closely spaced. If sufficient filtering is not provided at the output stages and/or on one or both of the transmitter feedlines, energy from one transmitter may travel back to the final power amplifier of the other. In the PA stage, the desired and undesired signals mix, producing spurious intermodulation products.

Consider the following example involving two transmitters. The third-order intermodulation products (IM3) are generated in the following manner:

- The output of transmitter 1 (f_1) is coupled into the nonlinear output stage of transmitter 2 (f_2). This condition is the result of insufficient isolation between the two output signals.
- Signal f_1 mixes with the second harmonic of f_2 , producing an in-band third-order term with a frequency of $2f_2 - f_1$ at the output of transmitter 2.
- In a similar manner, another third-order term will be produced at a frequency of $2f_1 - f_2$ at the output of transmitter 1.

¹ Portions of this section were contributed by Geoffrey N. Mendenhall, P. E., Cincinnati, OH.

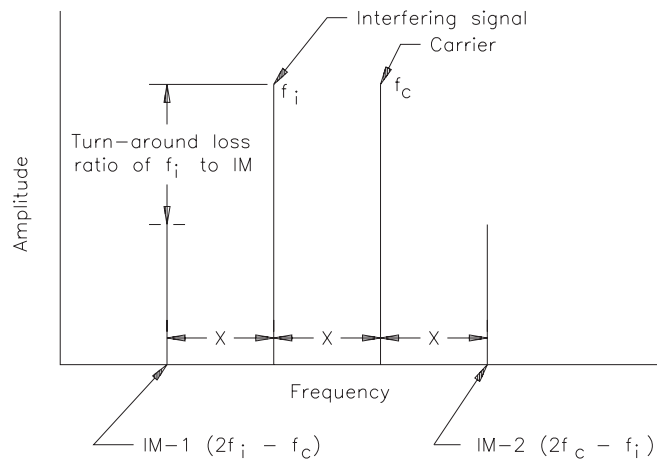


Figure 10.30 Frequency spectrum of third-order IM with the interfering level equal to the carrier level.

This implies that the second-harmonic content within the output stage of each transmitter, along with the specific nonlinear characteristics of the output stage, will have an effect on the value of the mixing loss.

It is also possible to generate these same third-order terms in another way. If the difference frequency between the two transmitters ($f_2 - f_1$), which is an out-of-band frequency, remixes with either f_1 or f_2 , the same third-order intermodulation frequencies will be produced.

Empirical measurements [3] indicate that the $2f_2 - f_1$ type of mechanism is the dominant mode generating third-order IM products in RF generators using a tuned cavity for the output network. Figure 10.30 illustrates the mechanism of third-order IM for the two-signal condition.

IM As a Function of Turnaround Loss²

Turnaround or *mixing loss* describes the phenomenon whereby the interfering signal mixes with the fundamental and its harmonics within the nonlinear output device. This mixing occurs with a net conversion loss; hence, the term *turnaround loss* has become widely used to quantify the ratio of the interfering level to the resulting IM3 level. A turn around loss of 10 dB means that the IM3 product fed back to the antenna system will be 10 dB below the interfering signal fed into the transmitter output stage.

Turnaround loss will increase if the interfering signal falls outside the passband of the transmitter output circuit, varying with frequency separation of the desired signal

2 Mendenhall, Geoffrey N., "A Study of RF Intermodulation Between FM Broadcast Transmitters Sharing Filterplexed or Colocated Antenna Systems, Broadcast Electronics, Quincy, IL, 1983.

and the interfering signal. This is because the interfering signal is first attenuated by the output circuit selectivity going into the nonlinear device, then the IM3 product is further attenuated as it comes back through the frequency-selective circuit. Turnaround loss, thus, can be divided into three principal components:

- The basic in-band conversion loss of the nonlinear device
- Attenuation of the out-of-band interfering signal resulting from the selectivity of the output stage
- Attenuation of the resulting out-of-band IM3 products resulting from the selectivity of the output circuit

As the turnaround loss increases, the level of undesirable intermodulation products is reduced, and the amount of isolation required between the transmitters also is reduced.

Transmitter output circuit loading directly affects the power amplifier source impedance and, therefore, also affects the efficiency of coupling the interfering signal into the output circuit where it mixes with the other frequencies present to yield IM3 products. Light loading reduces the amount of interference that enters the output circuit with a resulting increase in turnaround loss. In addition, output loading will change the output circuit bandwidth (loaded Q) and, therefore, also affect the amount of attenuation that out-of-band signals will encounter in passing into and out of the output circuit.

Second-harmonic traps or low-pass filters in the transmission line of either transmitter have little effect on the generation of IM products. This is because the harmonic content of the interfering signal entering the output circuit of the transmitter has much less effect on IM3 generation than the harmonic content within the nonlinear device itself. The interfering signal and the resulting IM3 products thus fall within the passband of the low-pass filters and outside the reject band of the second-harmonic traps; these devices typically offer little, if any, attenuation to RF intermodulation products.

Site-related intermodulation problems and typical solutions are discussed later in this chapter.

10.3.7 VSWR

Among all the possible load measurements, VSWR is the most common. An in-line directional wattmeter commonly is used to measure both the forward and reflected power in a feedline. It is the comparison of reflected power (P_r) to forward power (P_f) that provides a qualitative measure of the match between a transmitter and its feedline/load.

The standing wave ratio (SWR) on a transmission line is related to the ratio of the forward to reflected power by the following:

$$\text{SWR} = \frac{1 + \sqrt{P_r/P_f}}{1 - \sqrt{P_r/P_f}} \quad (10.14)$$

The standing wave is the result of the presence of two components of power: one traveling toward the load and the other reflected by a load mismatch, traveling back toward the generator. These components are defined by the following:

$$P_f = \frac{E_f^2}{Z_o} = I_f^2 (Z_o) \quad (10.15)$$

$$P_r = \frac{E_r^2}{Z_o} = I_r^2 (Z_o) \quad (10.16)$$

$$P_n = (P_f - P_r) \quad (10.17)$$

Where:

P_f = forward power

P_r = reflected power

E_f = forward voltage

E_r = reflected voltage

I_f = forward current

I_r = reflected current

Z_o = characteristic impedance of the line

P_n = net power absorbed by the load (transmission line loss and antenna radiation)

Because the forward and reflected voltages and currents are traveling in opposite directions, they will add in phase at some point along the line to produce a voltage maximum. At wavelength along the line in either direction from this maximum, the forward and reflected components will be out of phase, and will produce a voltage minimum. The forward and reflected components of current also add vectorally to produce a current standing wave. The magnitude of this wave is defined by:

$$\text{SWR} = \frac{E_{\max}}{E_{\min}} = \frac{I_{\max}}{I_{\min}} \quad (10.18)$$

The ratio of the highest voltage point on the line to the lowest voltage point on the line commonly is used to evaluate system performance. This property is known as the *voltage standing wave ratio* (VSWR). A VSWR of 1.0:1.0 means that a perfect match has been achieved. A VSWR of 2.0:1.0 means that a mismatch is causing power to be reflected back from the load, approximately 11 percent in this particular case.

At the point of reflection (the load mismatch), the phase of the reflected current is reversed 180° from the forward current. The reflected voltage does not exhibit this phase reversal. This displaces the voltage and current standing waves by 90° along the line so that the E_{max} and I_{min} occur at the same points, while E_{min} and I_{max} occur 90° in either direction from E_{max} and I_{min} .

The fact that the reflected current is reversed in phase makes it possible to measure forward and reflected power separately using a directional coupler device. A small voltage is obtained by inductive coupling that represents the current in the transmission line. To this signal is added a sample of the voltage across the line, simultaneously obtained by capacitive coupling. These two samples are adjusted to be exactly equal when the line is terminated with its characteristic impedance. The two RF samples, when added, yield a resultant RF voltage proportional to the forward components of voltage and current. Reflected power is measured by installing a second coupling section physically turned in the opposite direction. The phase of the current sample is reversed, and the reflected components add while the forward components balance out.

On a theoretical, lossless feedline, VSWR is the same at all points along the line. Actually, all feedlines have some attenuation distributed uniformly along their entire lengths. Any attenuation between the reflection point and the source reduces VSWR at the source. This concept is illustrated in [Figure 10.31](#). Standing waves on the feedline increase feedline attenuation to a level greater than normal attenuation under matched-line conditions. The amount of feedline attenuation increase depends on the VSWR and the extent of losses under matched-line conditions. The higher the matched-line attenuation, the greater the additional attenuation caused by a given VSWR. [Figure 10.32](#) correlates additional attenuation with matched-line attenuation for various VSWR levels.

When a feedline is terminated by an impedance equal to the feedline *characteristic impedance*, the feedline impedance at any point is equal to the feedline characteristic impedance. When the feedline is terminated by an impedance not equal to the feedline characteristic impedance, the feedline input impedance will vary with changes in feedline length. The greater the impedance mismatch at the load, the wider the feedline input impedance variation with changes in line length. The largest feedline impedance change occurs with the worst mismatches: open circuit or short circuit.

A short circuit at the end of a feedline appears as an open circuit at a point wavelength down the feedline. It is the greatest impedance transformation that can occur on a feedline. Less severe mismatches cause smaller impedance transformations. At a point wavelength down the line, the impedance equals the load impedance. Feedlines with high VSWR sometimes are called *resonant* or *tuned feedlines*. Matched lines—those with little VSWR—are referred to as *nonresonant* or *untuned feedlines*. This phenomenon means that high feedline VSWR can cause the feedline input to present a load impedance to the output tube that differs greatly from the feedline characteristic impedance. Some RF generators incorporate an output matching network that provides a *conjugate match* to a broad range of impedances.

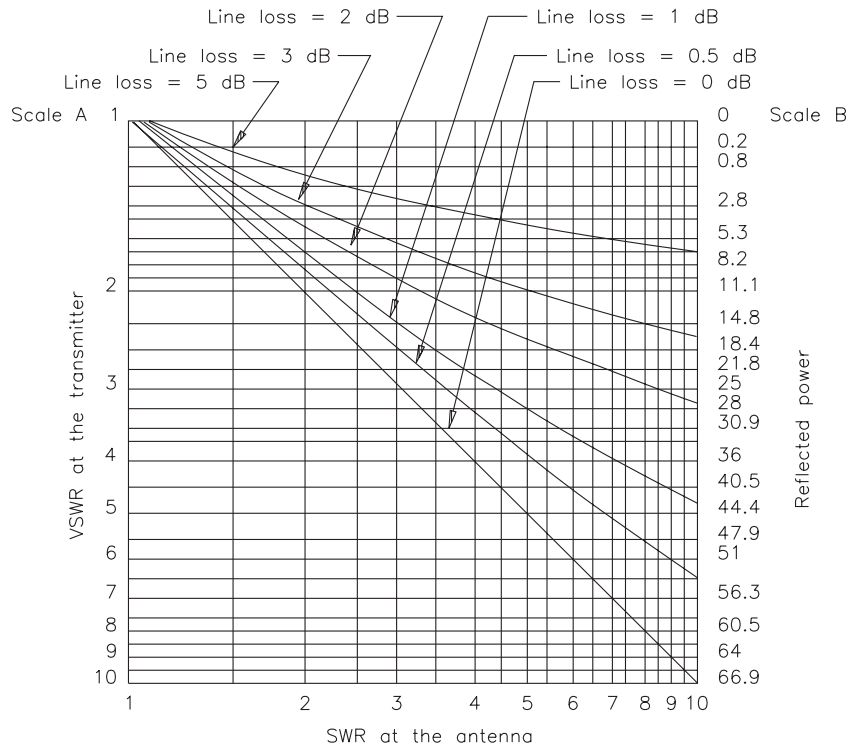


Figure 10.31 Illustration of how the VSWR on a transmission line affects the attenuation performance of the line. Scale *A* represents the VSWR at the transmitter; scale *B* represents the corresponding reflected power in watts for an example system.

10.4 RF System Performance

Assessing the performance of an RF generator has been greatly simplified by recent advancements in system design. The overall performance of most transmitters can be assessed by measuring the audio, video, or data performance of the system as a whole. If these “intelligence signal” tests indicate good performance, chances are good that all of the RF parameters also meet specifications. Both AM and FM transmission systems have certain limitations that prevent them from ever being a completely transparent medium. Through proper system design and regular equipment maintenance, however, overall performance can be maintained at acceptable levels in most operational environments.

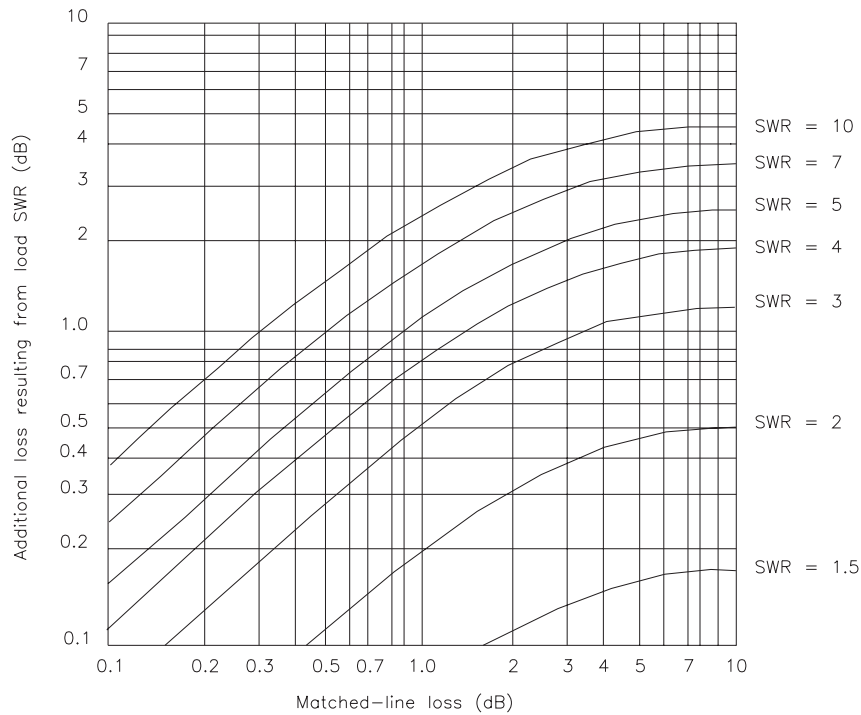


Figure 10.32 Illustration of how VSWR causes additional loss in a length of feedline.

10.4.1 Key System Measurements

The procedures for measuring transmitter performance vary widely, depending on the type of equipment being used. Certain basic measurements, however, can be applied to characterize the overall performance of most systems:

- Frequency response: the actual deviation from a constant amplitude across a given span of frequencies.
- Total harmonic distortion: the creation, by a nonlinear device, of spurious signals harmonically related to the applied waveform. THD is sensitive to the noise floor of the system under test. If the system has a signal-to-noise ratio of 60 dB, the distortion analyzer's best possible reading will be higher than 0.1 percent ($60 \text{ dB} = 0.001 = 0.1 \text{ percent}$).
- Intermodulation distortion: the creation, by a nonlinear device, of spurious signals not harmonically related to the audio waveform. These distortion components are *sum-and-difference* (beat note) mixing products. The IM measurement is relatively impervious to the noise floor of the system under test.

- Signal-to-noise ratio (S/N): the amplitude difference, expressed in decibels, between a reference-level audio signal and the residual noise and hum of the overall system.
- Crosstalk: the amplitude difference, expressed in decibels, between two or more channels of a multichannel transmission system.

10.4.2 Synchronous AM in FM Systems³

The output spectrum of an FM transmitter contains many sideband frequency components, theoretically an infinite number. The practical considerations of transmitter design and frequency allocation make it necessary to restrict the bandwidth of FM signals. Bandwidth restriction brings with it a number of undesirable side effects, including:

- Phase shifts through the transmission chain
- Generation of *synchronous AM* components
- Distortion in the demodulated output of certain receivers

In most medium- and high-power FM transmitters, the primary out-of-band filtering is performed in the output cavity of the final stage. Previous stages in the transmitter (the exciter and IPA) are designed to be broadband, or at least more broadband than the PA. As the bandwidth of an FM transmission system is reduced, synchronous amplitude modulation increases for a given carrier deviation (modulation). Synchronous AM is produced as tuned circuits with finite bandwidth are swept by the frequency of modulation. The amount of synchronous AM generated is dependent on tuning, which determines (to a large extent) the bandwidth of the system. [Figure 10.33](#) illustrates how synchronous AM is generated through modulation of the FM carrier.

Bandwidth Limiting

The design goal of most medium- and high-power FM transmission equipment is to limit the bandwidth of the transmitted RF signal at one stage only. In this way, control over system bandwidth can be tightly maintained, and the tradeoffs required in any practical FM system can be optimized. If, on the other hand, more than one narrowband stage exists within the transmitter, adjustment for peak efficiency and performance can be a difficult proposition. The following factors can affect the bandwidth of an FM system:

- Total number of tuned circuits in the system

³ Mendenhall, Geoffrey N., "Techniques for Measuring Synchronous AM Noise in FM Transmitters," Broadcast Electronics, Quincy, IL, 1988.

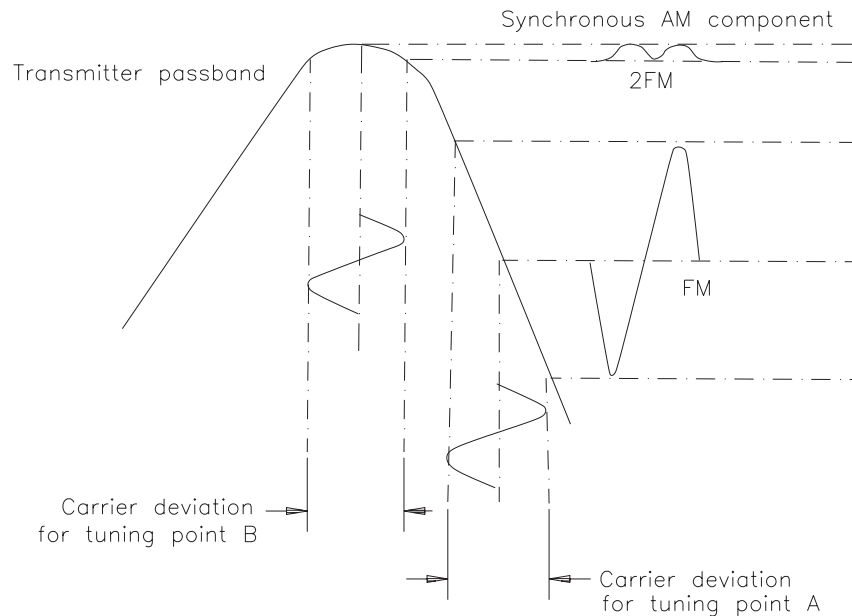


Figure 10.33 The generation of synchronous AM in a bandwidth-limited FM system. Note that minimum synchronous AM occurs when the system operates in the center of its passband.

- Amplitude and phase response of the total combination of all tuned circuits in the RF path
- Amount of drive to each RF stage (saturation effects)
- Nonlinear transfer function within each RF stage

The following techniques can be used to improve the bandwidth of a given system:

- Maintain wide bandwidth until the final RF stage. This can be accomplished through the use of a broadband exciter and broadband (usually solid-state) IPA.
- Minimize the number of interactive-tuned networks within the system. Through proper matching network design, greater bandwidth and simplified tuning can be accomplished.

Bandwidth affects the gain and efficiency of any FM RF amplifier. The bandwidth is determined by the load resistance across the tuned circuit and the output or input capacitance of the RF stage. Because grid-driven PA tubes typically exhibit high input capacitance, the PA input circuit generally has the greatest effect on bandwidth limiting in an FM transmitter. The grid circuit can be broadened through resistive swamping in the input circuit, or through the use of a broadband input impedance matching network. One approach involves a combination of series inductor and shunt capacitor elements

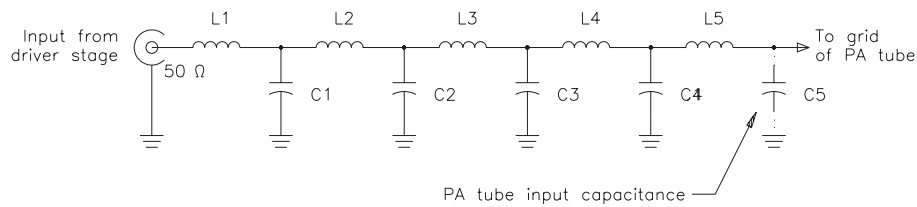


Figure 10.34 IPA plate-to-PA-grid matching network designed for wide bandwidth operation.

implemented on a printed wiring board, as illustrated in [Figure 10.34](#). In this design, the inductors and capacitors are etched into the copper-clad laminate of the PWB. This approach provides the necessary impedance transformation from a $50\ \Omega$ solid-state driver to a high-impedance PA grid in a series of small steps, rather than more conventional *L*, *Pi*, or *Tee* matching networks.

Effects of Synchronous AM

The effect of bandwidth on synchronous AM performance for an example FM RF amplifier is plotted in [Figure 10.35](#). Notice that as the $-3\ \text{dB}$ points of the RF system passband are narrowed, a dramatic increase in synchronous AM occurs. The effect shown in the figure is applicable to any bandwidth-limited FM system.

Bandwidth also affects the distortion floor of the demodulated signal of an FM system, as plotted in [Figure 10.36](#). The data shown apply to a test setup involving a single-tuned circuit fed by an FM generator with a deviation of $\pm 75\ \text{kHz}$ at a modulating frequency of $15\ \text{kHz}$ (no de-emphasis is applied to the output signal). Remember that FM is a nonlinear process and that interpretation of distortion numbers for multiple-carrier signals is not accurate. The example shown, however, illustrates how the bandwidth of the RF channel can set a minimum performance limit on system total harmonic distortion. By understanding the mechanics of synchronous AM, adjustment for optimum performance of an FM system can be accomplished.

10.4.3 Incidental Phase Modulation

Incidental phase modulation (IPM) is phase modulation produced by an AM transmitter as a result of amplitude modulation. In other words, as an AM transmitter develops the amplitude-modulated signal, it also produces a phase-modulated, or PM, version of the input intelligence. In theory, IPM is of little consequence for classic AM applications because FM and PM do not affect the carrier amplitude. An envelope detector will, in theory at least, ignore IPM. However, a number of applications utilize manipulation of the carrier phase to convey information in addition to amplitude modulation of the carrier. Receivers designed to decode such AM/PM waveforms will not perform properly if excessive IPM is produced by the transmitter. Al-

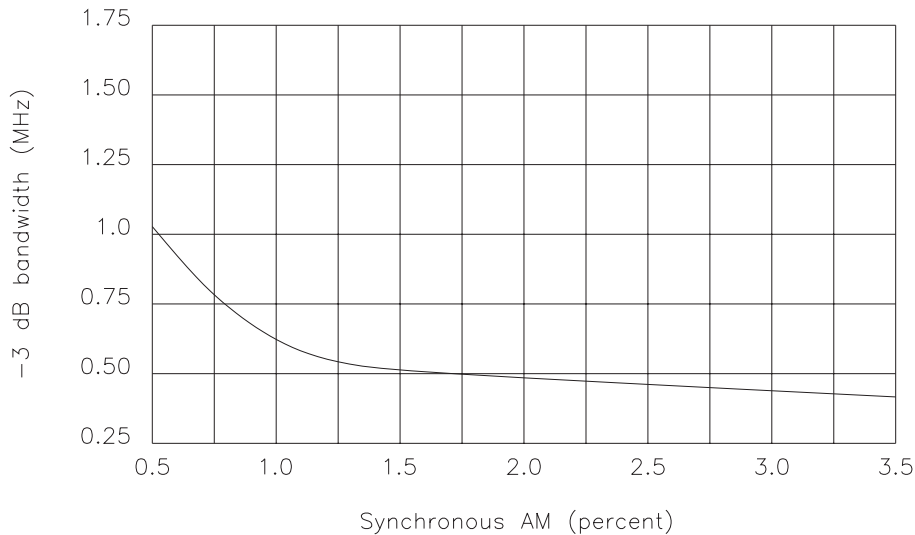


Figure 10.35 Synchronous AM content of an example FM system as a function of system bandwidth.

though the effects of IPM vary from one system to another, optimum performance is realized when IPM is minimized.

As a general rule, because IPM is a direct result of the modulation process, it can be generated in any stage that is influenced by modulation. The most common cause of IPM in plate-modulated and pulse-modulated transmitters is improper neutralization of the final RF amplifier. Adjusting the transmitter for minimum IPM is an accurate way of achieving proper neutralization. The reverse is not always true, however, because some neutralization methods will not necessarily result in the lowest IPM.

Improper tuning of the IPA stage is the second most common cause of IPM. As modulation changes the driver loading to the PA grid, the driver output also may vary. The circuits that feed the driver stage usually are sufficiently isolated from the PA that they do not produce IPM. An exception is if the power supply for the IPA is influenced by modulation. Such a problem could be caused by a lack of adequate capacitance in the filters of the high-voltage power supply.

Bandwidth Considerations

When a perfect AM transmitter is modulated with a tone, two sidebands are created, one above the carrier and one below it. Phase modulation of the same carrier, however, produces an infinite number of sidebands at intervals above and below the carrier. The number of significant sidebands depends on the magnitude of the modulation. With moderate amounts of IPM, the transmitted spectrum can be increased to the point at which the legal channel limits are exceeded.

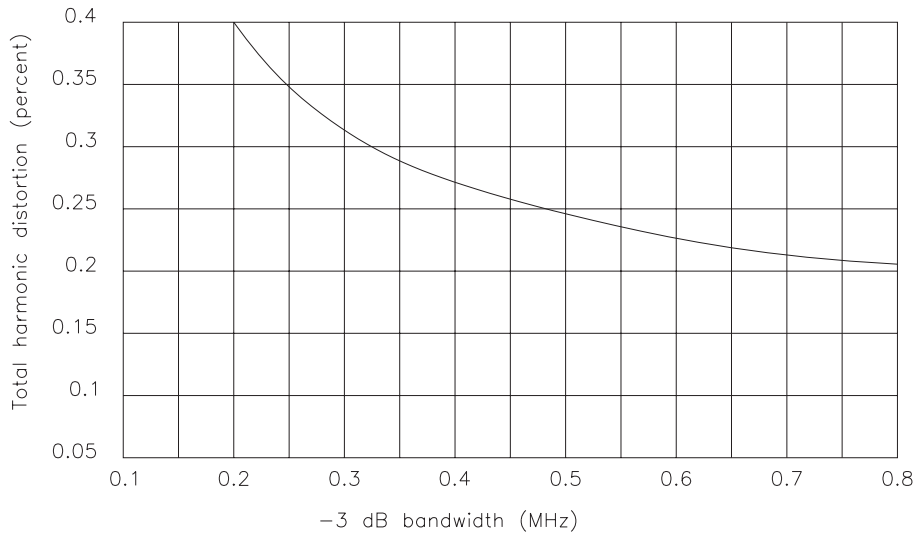


Figure 10.36 Total harmonic distortion (THD) content of an example demodulated FM signal as a function of transmission-channel bandwidth.

Corrective Procedures

The spectrum analyzer is generally the best instrument for identifying IPM components. The transmitter is first modulated with a pure sine wave, and a THD measurement is taken with a distortion analyzer. These data will provide an idea of what the spectrum analyzer display should look like when it is connected to the transmitter. If the transmitter has an excessive amount of harmonic distortion, it will be impossible to tell the difference between the normal distortion sidebands and the IPM sidebands.

Figure 10.37 shows the trace of an ideal transmitted spectrum on a spectrum analyzer when modulated with a 1 kHz tone at 100 percent modulation. Notice that there are just three components: one carrier and two sidebands. This transmitter would measure 0 percent distortion, because no harmonic sidebands are present. Also, no IPM is present, because it would show up as extra sidebands in the spectrum.

If the transmitter had an IPM problem, the spectrum analyzer display might look like Figure 10.38. The distortion analyzer still would read 0 percent, but the sidebands on the spectrum analyzer would not agree. The key is to start with a low harmonic distortion reading on the transmitter so that any sidebands will be recognizable immediately as the result of IPM.

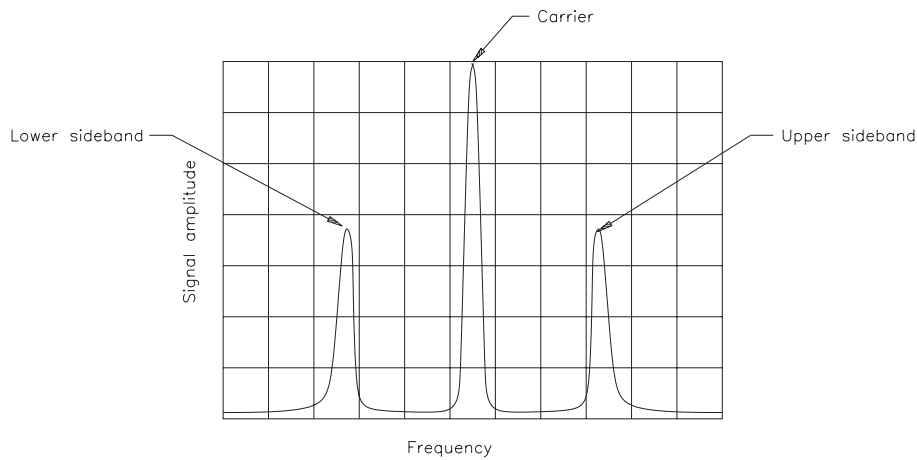


Figure 10.37 The spectrum of an ideal AM transmitter. THD measures 0 percent; there are no harmonic sidebands.

10.4.4 Carrier Amplitude Regulation

Carrier amplitude regulation (*carrier level shift*) is the effective shift in apparent carrier level as a result of the amplitude modulation process [4]. Carrier level shift can be caused by the following:

- Poor power supply regulation
- Modulation-related even order harmonic distortion that generates a dc offset component in the modulated RF envelope
- A combination of the previous two mechanisms

Carrier level shift can be either positive or negative, although is usually negative because power supply regulation is most often the major source of carrier level shift, and power supply regulation is generally negative in sign (i.e., lower voltage output at higher current loads). Strictly speaking, carrier level shift is defined as the shift in effective carrier voltage or current resulting from modulation. (See [Figure 10.39](#).)

10.4.5 Site-Related Intermodulation Products

As discussed previously in this chapter, an intermodulation product (intermod) is a signal created when two or more signals are unintentionally mixed to produce a new signal or group of signals. Mixing can occur in transmitter circuitry, in receiver circuitry, or elsewhere. The apparent effect is the same: one or more low-level products strong enough to interfere with desired signals and prevent proper reception. Some-

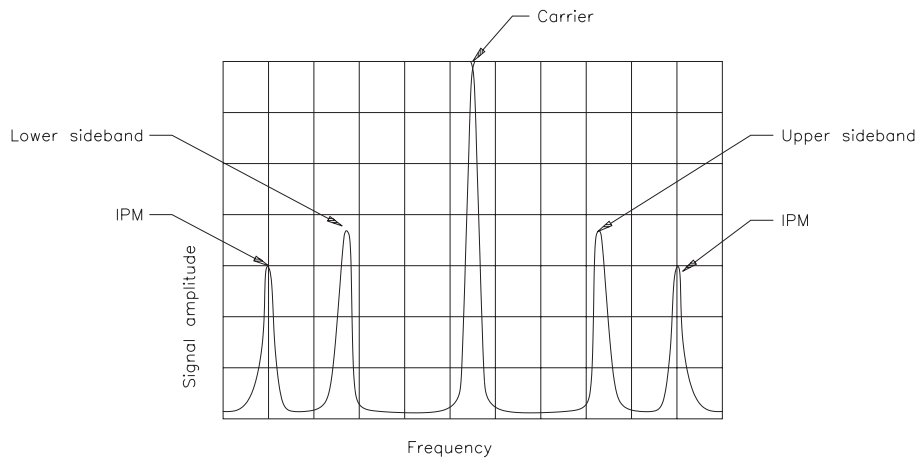


Figure 10.38 The spectrum of an ideal transmitter with the addition of IPM.

times the intermod falls on unused frequencies. Sometimes it falls on a channel in use or near enough to the channel to cause interference.

It is possible to calculate the intermod that may result from frequencies in use at a given transmission site. Normally produced by a computerized analysis, this *intermod study* is helpful in several ways. If the study is conducted as one or more new frequencies are under consideration for future use at an existing site, it may reveal which frequencies are more likely to produce a harmful intermod and which are not. If the study is conducted as part of an investigation into an interference problem, it may help in identifying transmitters that play a role in generating the intermod.

Intermod can be generated by any nonlinear device or circuit. The most common place for intermod to be generated is in the power amplifier of an FM transmitter. Such an amplifier does not need to preserve waveform integrity; the waveform is restored by the tuned output circuit. Thus, the amplifier is operated in an efficient but nonlinear mode, class C. In contrast, many amplifiers for AM service operate in a linear mode. Such amplifiers normally do not act as mixers.

Generally, solid-state amplifiers are better mixers than their electron tube counterparts, a fact that may explain why intermod may appear suddenly when older equipment is replaced. A mix occurs when RF energy from one transmitter (Tx 1) finds its way down the transmission line and into the final amplifier of another transmitter (Tx 2), as illustrated in [Figure 10.40](#). The energy mixes in the final amplifier with the prefiltered second harmonic of Tx 2 to produce another in-band signal, a third-order intermod product. The reverse mix also may occur in the final amplifier of Tx 1, producing another intermod signal.

After the cause is discovered, a mix generated in a transmitter final amplifier is relatively easy to eliminate: Increase the isolation between the offending transmitters. This

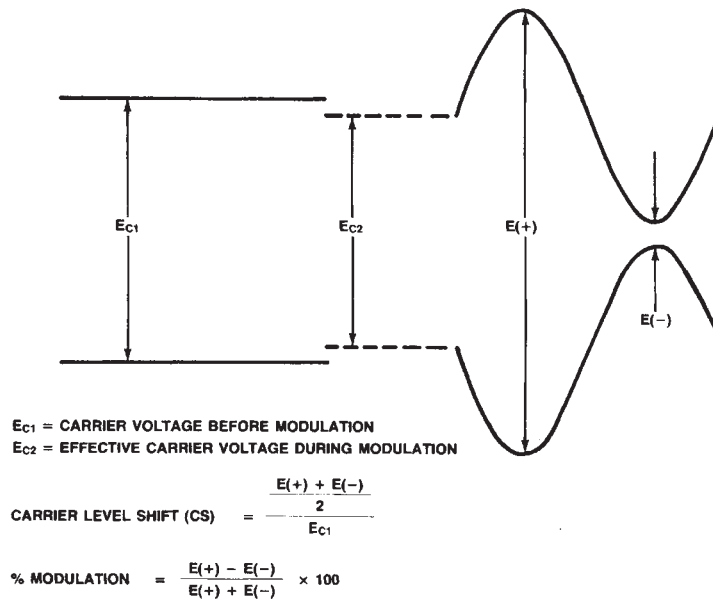


Figure 10.39 Carrier amplitude regulation (carrier level shift) defined. (From [4]. Used with permission.)

may be accomplished by increasing the space between the antennas, or by placing an isolator at the output of each transmitter.

Receiver Intermod

Intermod also may be generated in the front end of a receiver. High-level RF energy from transmitters at a site can generate a bias voltage in the receiver sufficient to place the receiver RF amplifier into a nonlinear state. In such a mode, the stage acts as a mixer and creates unwanted frequencies. Filters may be used to solve this problem. The filters should notch out the offending transmitted signals. If another transmitter is added to the site, another notch filter usually must be added to the receiver transmission line.

Hardware Intermod

At a large transmission site, the entire area—including towers, guy wires, antennas, clamps, and obstruction lighting hardware—is saturated by high-level RF. Any loose or moving hardware can cause broadband noise. Corroded connections, tower joints, or clamps can form semiconductor junctions that may generate harmful intermod. Because the antenna system is exposed to a changing external environment, the de-

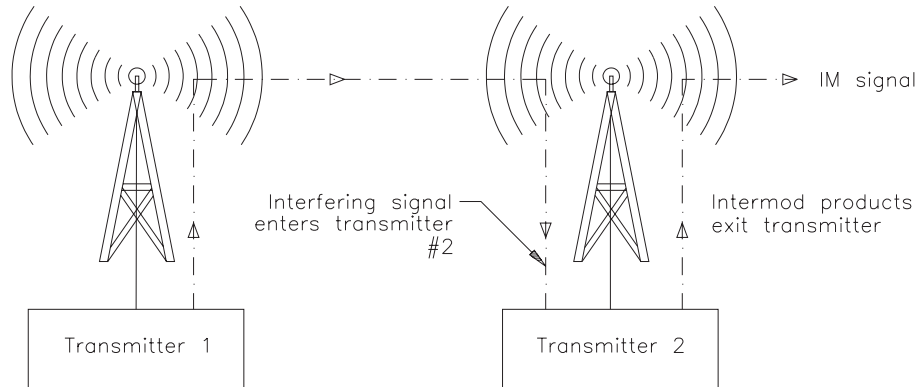


Figure 10.40 The mechanics of intermod generation at a two-transmitter installation.

gree of semiconductivity can vary with moisture or temperature. Changing semiconductivity causes intermittent problems, at changing levels of severity.

10.5 References

1. Hulick, Timothy P., "Very Simple Out-of-Band IMD Correctors for Adjacent Channel NTSC/DTV Transmitters," *Digital Television '98*, Intertec Publishing, Overland Park, KS, 1998.
2. Laboratory Staff, *Care and Feeding of Power Grid Tubes*, Varian/Eimac, San Carlos, CA, 1984.
3. Mendenhall, Geoffrey N., M. Shrestha, and E. Anthony, "FM Broadcast Transmitters," *NAB Engineering Handbook*, 8th Ed., E. B. Crutchfield (ed.), National Association of Broadcasters, Washington, DC, pp. 463 - 464, 1992.
4. George Woodard, "AM Transmitters," in *NAB Engineering Handbook*, 9th Ed., Jerry C. Whitaker (ed.), National Association of Broadcasters, Washington, DC, 1998.

10.6 Bibliography

- Benson, K. B., and Jerry C. Whitaker, *Television and Audio Handbook for Engineers and Technicians*, McGraw-Hill, New York, 1989.
- Bordonaro, Dominic, "Reducing IPM in AM Transmitters," *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1988.
- Crutchfield, E. B. (ed.), *NAB Engineering Handbook*, 8th Ed., National Association of Broadcasters, Washington, DC, 1992.

- Dolgosh, Al, and Ron Jakubowski, "Controlling Interference at Communications Sites," *Mobile Radio Technology*, Intertec Publishing, Overland Park, KS, February 1990.
- Hershberger, David, and Robert Weirather, "Amplitude Bandwidth, Phase Bandwidth, Incidental AM, and Saturation Characteristics of Power Tube Cavity Amplifiers for FM," application note, Harris Corporation, Quincy, IL, 1982.
- Kinley, Harold, "Base Station Feedline Performance Testing," *Mobile Radio Technology*, Intertec Publishing, Overland Park, KS, April 1989.
- Mendenhall, Geoffrey N., "A Study of RF Intermodulation Between FM Broadcast Transmitters Sharing Filterplexed or Co-located Antenna Systems," *Broadcast Electronics*, Quincy, IL, 1983.
- Mendenhall, Geoffrey N., "A Systems Approach to Improving Subcarrier Performance," *Proceedings of the 1986 NAB Engineering Conference*, National Association of Broadcasters, Washington, DC, 1986.
- Mendenhall, Geoffrey N., "Fine Tuning FM Final Stages," *Broadcast Engineering*, Intertec Publishing, Overland Park, KS, May 1987.
- Mendenhall, Geoffrey N., "Improving FM Modulation Performance by Tuning for Symmetrical Group Delay," *Broadcast Electronics*, Quincy IL, 1991.
- Mendenhall, Geoffrey N., "Techniques for Measuring Synchronous AM Noise in FM Transmitters," *Broadcast Electronics*, Quincy IL, 1988.
- Mendenhall, Geoffrey N., "Testing Television Transmission Systems for Multichannel Sound Compatibility," *Broadcast Electronics*, Quincy, IL, 1985.
- Terman, F. E., *Radio Engineering*, 3rd Ed., McGraw-Hill, New York, 1947.
- Whitaker, Jerry C., *Maintaining Electronic Systems*, CRC Press, Boca Raton, FL, 1992.
- Whitaker, Jerry C., *Radio Frequency Transmission Systems: Design and Operation*, McGraw-Hill, New York, 1990.
- Witte, Robert A., "Distortion Measurements Using a Spectrum Analyzer," *RF Design*, Cardiff Publishing, Denver, CO, pp. 75 - 84, September 1992.